THE
RICE INSTITUTE PAMPHLET

Vol. XVI October, 1929 No. 4

Published by
THE RICE INSTITUTE

A university of liberal and technical learning founded by William Marsh Rice in the City of Houston, Texas, and dedicated by him to the advancement of Letters, Science, and Art

THE
RICE INSTITUTE PAMPHLET

Vol. XVI October, 1929 No. 4

Published by
THE RICE INSTITUTE

A university of liberal and technical learning founded by William Marsh Rice in the City of Houston, Texas, and dedicated by him to the advancement of Letters, Science, and Art

THE RICE INSTITUTE PAMPHLET
Vol. XVI OCTOBER, 1929 No. 4

CONTENTS

PAGE

THE NATURE OF MIND—Three public lectures delivered in the Physics Amphitheatre of the Rice Institute, January 6, 13, and 20, 1929, by CHARLES WILLIAM MORRIS, Ph.D. (Chicago), Assistant Professor of Philosophy at the Rice Institute.

I. The Problem of Mind and Its Historical Setting 153
II. Three Theories of Mind 183
III. The Functional View of Mind 214

MATHEMATICAL LECTURES delivered at the Rice Institute, May 20, 22, and 23, 1929, by HERMANN WEYL, Ph.D. (Göttingen), Professor of Mathematics at the Technische Hochschule in Zürich.

I. Consistency in Mathematics. 245
II. The Spherical Symmetry of Atoms 266
III. Gravitation and the Electron 280

THE NATURE OF MIND[1]

I

THE PROBLEM OF MIND AND ITS HISTORICAL SETTING

MOST of us would defend to the last ditch the existence and worth of our own minds, but we easily become embarrassed if we are asked to say very much about the nature of that mind which we so hotly defend. And it is no answer to reply "never matter" for in physical terms it is doubtful if mind and matter are so opposed as we ordinarily believe, and in social terms, it is certain that nothing matters more than mind.

In raising the problem of the nature of mind we are plunged into a problem of the greatest difficulty and of the deepest importance. To be sure, many thinkers, like John B. Watson, have called mind a "passing concept"; nevertheless, the abolition of a name gives no solution, but merely forces old problems to appear under new labels and in different disguises. The term "mind" occupies too great a place in the history of human speculation to be lightly dismissed. It is more courageous to confront the problem boldly, and to inquire into what has been meant, and what should be meant, by the term mind. Although it is impossible to say all that should be said, the aim at least must be to suggest a precise answer to this problem. In the present lecture the historical aspects of the problem will be raised; in the second lecture three typical contemporary theories of mind will be considered; and in the final lecture the attempt will be made to defend a fourth type of theory and to develop some of its relevant implications.

---

[1] A series of three lectures delivered at the Rice Institute on January 6, 13, and 20, 1929, by Charles William Morris, Ph.D., Assistant Professor of Philosophy at the Rice Institute.

I

The stream of contemporary reflective thought appears to have taken one dominant direction, that of assuming a critical attitude to thought itself, of subjecting the most fundamental concepts to a merciless dissection. If one has an ear for subtle music, there is to be heard behind the philosophical display the ever repeated note of an impatience with thought and words, as though thought and words somehow ensnared one and obstructed participation in the fascinating drama which is felt to lie on beyond. "I'm bored, brother, with human words—all our words. Bored! I've heard each one of them a thousand times surely," sighs a Gorki character in *The Lower Depths,* and in the "higher depths" of James and Bergson the lament is repeated: James "abandoning" logic to plunge into "the stream of life" whose "deeper features of reality are found only in perceptual experience"; Bergson wishing to "reabsorb intellect in intuition," for intuition leads "to the very inwardness of life" while intellect "goes all around life" without ever penetrating into life. Better to be afloat in a sea of chaotic purposelessness—so runs many a modern tale—than to be enmeshed in, or to see the world enmeshed in, a logical or verbal net.

The resulting opposition to rationalism actually depreciates the intellect in some cases, but in general the modern effort has been to put intellect back into its larger setting, without neglecting its importance within this setting. Accordingly, it has been widely recognized that a basic task of critical philosophy, to use Broad's term, is the clarification of the fundamental concepts of reflective discourse.

Such an inquiry into basic concepts constitutes one significant aspect of the remarkably critical spirit which is penetrating contemporary science. It has been increasingly realized that a great deal of nonsense can be covered by the magic word "science." The boldness of scientific advance has gone hand in hand with a scrutiny of the basic concepts. This is as true of the behavioristic movement in psychology with its reaction against the uncritical use of such terms as mind and consciousness, as it is of the physical theory of relativity with its attempt to use precisely such terms as simultaneity, length, and gravity. All such concepts have felt, and will continue to feel, the probing, burrowing knife-edge of analytical thought. Something is always likely to happen when a thinker asks himself what he means, in precise terms, by the fundamental concepts he employs.

In this analysis of concepts, the concept of mind occupies a peculiarly important place. Much, in fact, may be said in support of the claim that the problem of mind is the central problem of contemporary philosophy.

In the first place, the solution of the meaning of meaning —and if meaning is not mental, nothing is—is basic to the problem of how to determine any specific meaning. If there is to be a philosophy of science, the problem of the place of mind in scientific procedure must be determined. And since the topic of mind touches upon the work of the psychologist, biologist, anthropologist, and sociologist, the problem allows of that type of synthetic activity which certain philosophers have always regarded as one aspect of desirable philosophical activity. Philosophy and science are again drawing much closer together, and the topic of mind furnishes a fruitful point of coöperation.

Within philosophy itself, the central importance of an adequate theory of mind is much more obvious. The common classification of philosophers into realists and idealists centers around the importance attributed to mind. For the realist, mind is at best one aspect or portion of a larger non-mental universe which is not dependent upon mind for either its nature or existence; for the idealist mind is given a much greater metaphysical significance, the universe being conceived either as a mind or collection of minds, or as dependent upon mind both for its nature and existence. Mind occupies a central place in the problem of knowledge, which is usually stated as involving the relation of the mind or knower to the reality known. When so stated, it is clear that the form of the inquiry and its termination is largely dependent upon whatever attitude the inquirer takes to mind. In this sense, the problem of mind underlies the whole epistemological quest. Although the connection is not so intimate, the problem of mind bears upon the topics of logic and the philosophy of mathematics. If logic is to be the study of the entire reflective process, its ultimate ground must lie in a theory of mind, and its farthest reach must include the mathematical type of reflection.

Finally, in addition to the relation of the problem of mind to the problems of science, and to the basic problems of philosophy, there is the relation of mind to our individual and social life. We are still seeking for the essence of that rationality which Aristotle took to be the essence of man. In this psychozoic age, when from every quarter of the intellectual realm we are made to realize the supreme importance of reflective intelligence for morality, for education, and for internationalism, it would be superfluous to elaborate the advantages which a clear understanding of this priceless tool would afford.

The extended ramifications and implications of the topic of mind provide the only set of valid criteria for an adequate theory of mind. Any term can, of course, be used in any way the user wishes, and for special purposes of psychology and biology theories may be adequate which are not adequate in the domains of logic and metaphysics. A satisfactory theory of mind must be comprehensive in its scope and fertile in its implications. The following criteria may be proposed: the theory of mind must be neither so general nor so limited in character that it fails to clarify the philosophical discussions of mind; it must be in harmony with whatever material the various psychological points of view can offer; it must harmonize with the general theory of the development of animal intelligence; it must be applicable to the study of the mind of primitive man; it must throw light upon the relation of logic and mathematics to mental phenomena; and finally, it must not be what Broad calls a "silly theory," that is a theory which one hotly defends before his colleagues but would hesitate to mention to one of his friends of good common sense. And if these requirements form a staggering array, it can only be said that nothing less is acceptable. So much, then, for the standards which shall guide us in our quest. For the present, it will be well to survey briefly and in general outlines the historical development which the philosophy of mind has undergone.

II

If a large generalization is wanted, it can be said, in Hegelian fashion, that there have been three "moments" in the history of speculation concerning mind: a period in which mind and matter were only vaguely conceived and vaguely differentiated; a period in which mind and matter were sharply differentiated and opposed; a period, lasting up to the present, in which the effort has been to restore again the intimate connection of mind and matter by obtaining a new concept of nature. Although these periods overlap, and although traces of the earliest speculation continue to the present day and suggestions of the most recent theories can be found in the earliest period of speculation, it is roughly true that the history of thought about mind reveals a passage from the first, through the second, to the third of the above positions. For the remainder of this lecture we will trace the above development historically, reserving until later the critical analysis of the three main types of theory of mind which the historical process has produced.

As is to be expected, early man has no sharp separation of the life of mind from the rest of the world in which he dwells. Man only becomes conscious of himself and his reflective gropings as the result of a long historical development, in which by continual struggle he is forced to pay attention to himself and to the tool upon which his knowledge depends. Intelligence is normally directed outward; the science of man is always the last science to appear. Long before he knows himself man has made friends with the stars. The attitude of primitive man is therefore an objective attitude which allows of no opposition of mind and nature, of a private subjective life of consciousness and an outward objective world of corporeal events. For him there can be no question as to how mind and matter can interact, or how mind can know a world that is not mind. Similarly, at the beginning of human thought, there is no sharp distinction of a spiritual soul and a physical body. The "after-world" of early man is a world patterned after the lines of the terrestrial world; the future life is a glorified earth life, a continuation and not another kind of life. Death as conceived by the materialist is a concept of recent origin and utterly foreign to the early stages of thought. The situation that has been referred to is reflected in the poems of Homer, in the Rig Veda, and in the Old Testament of the Hebrews.

The same condition prevails among the early philosophers. Again there is no awareness of the problem of knowledge, no opposition of mind and the world. Mind, in fact, is not regarded as private and personal, but rather as *the principle of motion and order of the world*. The most distinctive aspect of the earliest theories of mind is precisely their lack of a psychological orientation. The early Logos or World-Reason is objective mind. Heraclitus[1] hints at a Reason which guides all things. Parmenides, in an often quoted statement that "it is the same thing that can be thought and that can be," hints vaguely at a doctrine that has been regarded as the prototype of the absolute idealisms. Empedocles[2] says of God, "He is only mind, sacred and ineffable mind, flashing through the whole universe with swift thoughts." The most interesting reference is perhaps the famous statement of Anaxagoras,[3] which is of historical importance because of its effect upon Plato: "mind is infinite and self-ruled and is mixed with nothing . . . it is the thinnest of all things and the purest, and it possesses all knowledge and the greatest power." "Over all is mind the ruler. And over the whole revolving universe mind held sway, so that it caused it to revolve in the beginning." "Mind knows all things . . . regulated all things." "All mind . . . is alike." "Mind . . . is eternal." In this statement there is evident the emotional glorification of mind, the absence of a sharp distinction between mind and nature, and the use of mind as the ultimate principle of motion.

---

[1] Bakewell, *Source Book in Ancient Philosophy*, p. 32, Fragment 72.

[2] Ibid., p. 46, Fragment 134.

[3] Ibid., pp. 51, 52, Fragment 12.

The existing fragments of early thought that can be called psychological show no satisfactory conception of the relation of thought to sense phenomena. Aristotle says that the atomists identified sense and reason.[1] Windleband says[2] of Empedocles that "the mixed substance which streams through the living body, the blood, was regarded as soul"; and that for the Sophists,[3] "the entire psychical life consisted only in perception." At the opposite pole, there is evidence in Parmenides, and perhaps in Democritus, of a sharp separation of thought and sense which led almost to the denial of the latter. Psychologically, then, the early thinkers either failed to discriminate between mental processes and the phenomena of sense, or indiscriminately opposed each to the other.

There thus arises in the ancient world a double problem of mind: metaphysically, the problem of the relation of mind as a principle of movement to the objects moved; psychologically, the problem of the relation of mind to the phenomena of sense. It cannot be said that even Plato or Aristotle solves either of these problems. Instead, the two problems merely become more acute.

It is Plato's merit to have become aware of the difficulties which universals present. Thought is concerned with whiteness, with man, with justice, with circles, and yet perception reveals only white things, individual men, just acts, and things approximately circular. It is clear, then, that thought and sense are not identical. But the superstructure which Plato builds upon this undeniable fact only raises new difficulties. Plato is first of all a moralist. Believing that if reality is only an ever-changing process, knowledge is impossible, and that morality depends upon knowledge, Plato became intoxicated with the implications for morality which the existence of universals entails. Instead of entering into an analytical discussion of how universals are related to particular things, he passes at once to the conclusion that there must exist a realm of objects which are not transitory, and which make possible the existence of genuine knowledge. These objects of knowledge constitute the realm of Platonic Ideas. It is impossible to enter at this point into the maze of controversy concerning the Platonic doctrine of Ideas, but without a great deal of reinterpretation it cannot be denied that these Ideas as used by Plato are not in any way mental. They are objects of mind, not objects in mind. This is a striking example of the non-psychological character of the Platonic analysis. It is necessary to look elsewhere for Plato's treatment of mind.

---

[1] *De Anima,* I, 404.

[2] Windleband, *A History of Philosophy,* p. 62.

[3] Ibid., p. 91.

As early as the sixth century before Christ, religious influences had led to the separation of the body from the soul,[1] and had set up rites for the soul's purification. In the *Phaedo* Plato thinks in terms of this same separation of soul and body. The body is the prison house of the soul, which attains its release only at death. "The soul is only able to view existence through the bars of a prison, and not in her own nature."[2] Accordingly the world of sense becomes a hindrance to knowledge and is sharply separated from thought: "thought is best when the mind is gathered into herself . . . when she has as little as possible to do with the body."[3] When the soul reflects "she passes into the realm of purity, and eternity, and immortality, and unchangeableness."[4]

---

[1] Burnet, *Early Greek Philosophy,* p. 86.

[2] *Phaedo,* 82.

[3] Ibid., 65.

[4] Ibid., 79.

Although the *Phaedo* exaggerates the dualistic aspect of Plato's thought, it seems just to say that the rupture of thought and sense which is found in this dialogue is not remedied elsewhere. Even in the passages which some have regarded as anticipations of behaviorism, such as the statement that "the soul when thinking appears to me to be just talking—asking questions of herself and answering them, affirming and denying,"[1] the emphasis is not so much on the relation of thinking to language, as on thinking as an activity of the soul. This is more evident in the similar passage in the *Sophist* where it is suggested that "thought is the unuttered conversation of the soul with herself."[2] Specifically, "the soul, besides other things, contains mind."[3] It is the immaterial soul that knows and thinks, and the sense world is at the best a stimulus to this activity.

Looked at metaphysically, the soul plays an important part in the cosmic process, as well as in the human body, as the principle of movement and order. It "holds and carries and gives life and motion to the entire nature of the body"[4]; it "is among the first of bodies and before them all, and is the chief author of their changes and transpositions."[5] The soul as the "self-moved" is "the nature which controls heaven and earth, and the whole circumference."[6] So closely are soul and mind related that Plato does not hesitate to refer to the "mind that orders the universe,"[7] and to call mind "the ruling power."[8] Plato's conception of the creation of the world is of a creation of order, not of stuff, a creation in accordance with the Ideas or Eternal Patterns, which remain, to the very end, objects and not contents of mind.[1]

---

[1] *Theaetetus*, 189, 190.

[2] *Sophist*, 263.

[3] *Laws*, 961.

[4] *Cratylus*, 400.

[5] *Laws*, 892.

[6] Ibid., 897.

[7] Ibid., 966.

[8] *Timaeus*, 48.

Plato's theory, in sum, does not solve, but rather accentuates the dualism of thought and sense, and brings into bold relief the doctrine of objective mind as the principle of motion and order in the universe.

If we are to characterize briefly Aristotle's relation to the Platonic treatment of mind, we may say that mind remains a principle of motion and order, but that this function is diminished by the larger place that is given to the Forms and Ideas as causal factors, and by an increased emphasis on the material or efficient cause. Aristotle also makes a valiant but ultimately unsuccessful attempt to bridge the Platonic gap between pure thought and sense. Although Aristotle's tendency is to oppose the substantive theory of mind in attempting to integrate mind and nature, the integration is never completed and mind remains a metaphysical outcast from nature. But these general statements must be given a more detailed and solid content.

For Aristotle, the world is an eternal process with no single goal, but always reaching specific consummations. The substratum which takes on specific forms he calls matter; specific things are what they are because of the Forms or Archetypes which guide the specific developmental processes. The Form of the oak tree guides the process of the developing acorn, and is the actuality which the acorn, as a potential oak tree, becomes. If to modern minds this analysis seems artificial, it must be remembered that Aristotle, lacking the causal knowledge which would be utilized today, is attempting to explain the fact of kinds or types in nature. What to us seems merely a verbal solution—the forms of things are due to the Forms—seemed to Aristotle a valid explanation.

---

[1] Ibid., 29, 32.

According to the more empirical phase of Aristotle's thought, Form and matter are correlative, the Forms are in matter, and matter is informed. Although the Forms are regarded as causes, they are not the only causes, and in recognizing the efficient or material cause, Aristotle is already beginning to limit the sphere in which immaterial causes are invoked to explain material processes.

The Forms of living beings Aristotle calls soul.[1] These Forms are somehow able to receive or grasp the Forms of other objects. This is the process of thought and mind, and in this explanation Aristotle is making mind merely one consummation of the process of nature. Mind has become the locus of Forms, "the Form of all Forms."[2] Further, there is no pure thought, no thought without content of nature present: "what we call reason in the soul . . . is, prior to the exercise of thought, no reality at all."[3] In fact, "the soul never thinks without the use of images."[4] So that while thinking differs from sense experience, which is the reception of the sensible patterns of things without their matter, thinking, the apprehension of Form, itself requires experiential content.

In this doctrine, whatever the difficulties in the concept of Form, it is clear that mind is, to use a contemporary phrase, an emergent aspect of the world process, and not a substance different in kind from this process. It would seem also, at first sight, that Aristotle has brought together the realm of the thought and the realm of experience. But both such conclusions would be premature, since Aristotle himself admits that there are some Forms that are not embodied in matter, namely, God, the intelligences of the spheres, and a part of the human soul which Aristotle calls "active reason." It may be well to glance briefly at the conceptions of God and active reason.

---

[1] *Metaphysics,* 1035b.

[2] *De Anima,* 429a, 432a. See also Ross's *Aristotle,* p. 132.

[3] *De Anima,* 429a.

[4] Ibid., 431a. Cf., 427b, 432a.

Aristotle's conception of Deity is of interest in connection with the topic of mind, because he conceives of Deity both as an unmoved mover, and as pure thought thinking itself. In the conception of an unmoved mover (as contrasted with Plato's self-moved mover), that is, of a principle of motion which directs the heavenly spheres in virtue of the urge or love of these spheres for perfection, we encounter again the use of the category of soul or mind as an explanation of motion. And in saying that "in all things the good is in the highest degree a principle,"[1] Aristotle has practically deserted his doctrine of specific teleologies in favor of the Platonic doctrine of a single universal purpose determined and directed by mind. Further, in conceiving God as thought thinking itself, Aristotle has again on his hands a metaphysically pure thought not thinking anything external to itself, not in any way dependent upon nature, and not requiring for its activity the slightest vestige of an image.

In fact, even in Aristotle's psychology of the finite mind, the diremption of thought and experience is inevitable. Under the term passive reason, Aristotle lumps together the phenomena of sensation, imagination, memory, and recollection.[2] The passive reason is treated empirically, and is the "matter" upon which the active reason works.[3] The distinction between the passive and active reason is clearly akin to the distinction between conception and perception (when this ambiguous term is equivalent to immediate experience). Instead of tracing empirically the process by which concepts arise in the world of experience, Aristotle deserts here the empirical point of view, and in line with his metaphysical doctrines, regards the active reason as that aspect of the soul which is not intrinsically connected with matter, but which somehow enters the biological process through the instrumentality of the semen,[1] and which is capable of existing apart as a sort of wax tablet which is as yet unwritten upon.[2] The result is that the higher phases of mind have become a substance separable from nature, and the conceptual aspect of thought does not develop in an unbroken line from the non-conceptual aspects of experience.

---

[1] *Metaphysics,* 1075a.

[2] Hammond, *Aristotle's Psychology,* LXXXIV.

[3] Ibid., LXXVIII: "Active reason stands to passive reason in the relation of form to matter.

Although Aristotle, like Plato, was justified in insisting that thought is not identical with the flux of experience, Aristotle as well as Plato turned to the opposite extreme of the Sophist in so emphasizing the uniqueness of thought that its relation to the experiential world remains an enigma. From its origin to its termination, the main course of Greek thought failed metaphysically to explain the relation of mind to nature, and failed psychologically to explain the relation of thought to experience. These two problems are handed down by the ancient thinkers to the thinkers of the modern world.

III

In passing to the modern period, it would be well to recall a commonly recognized fact that is of the greatest importance. With the increasing disintegration of the Greek culture, the individual was thrown back more and more upon his own resources, and became increasingly aware of himself as a being struggling for an adjustment which the existing social groups could no longer supply. The reconstructed society which Plato and Aristotle built in theory was not realized in actuality. While for Aristotle, if certain qualifications required by his metaphysics are omitted, the goal of man was to rise to his fullest stature within the community in which he participated, by the time of Plotinus the salvation on the earth had become transformed into a salvation from the earth. Even the Roman Stoics of the time of Christ, Seneca and Marcus Aurelius for example, frequently referred to the body as the prison house of the soul. Philosophy had become increasingly devoted to ethical and religious ends.

---

[1] *De Generatione Animalium,* 737a.

[2] *De Anima,* 430a.

The net result of this situation is that it favored a subjective interpretation of the categories of mind and consciousness, and made these categories the center of attention. The substance view of mind that is found in Plato (at least the view that mind is an attribute of the soul-substance) came to dominate philosophical reflection.

The separation of mind from nature is augmented during the Middle Ages. Saint Augustine's use of the "I doubt, therefore I am" argument shows how natural it had become to feel that mental phenomena furnished the sole starting point for rigorous thought. For him, the intellect is alone infallible.[1] Truth can be obtained only by the pure intellect, since in experience man is not in direct contact with the external world.[2] Although Thomas Aquinas temporarily stems the tide of subjectivity by an analysis of mind and knowledge essentially Aristotelian, nominalism and subjectivism continue to gain the ascendant, and in William of Occam there is encountered a situation strikingly similar to that of Descartes: mind is sharply separated from the processes of external nature, so that what is present is never the nonmental world itself. Knowledge is limited to the "signs" of things, and cannot deal with things directly.[1] With a supernaturalistic metaphysics, and an emphasis upon the inner life, it had become increasingly easy to think of mind as differing in essence from the external world, and to regard the qualitative world of experience as mental and as "in the mind."

---

[1] J. Martin, *Saint Augustin*, p. 36.

[2] Ibid., pp. 265, 277. Cf. Augustine's *De Diversis Quaestionibus*, Quaest. IX: "non est igitur exspectando sinceritas veritatis a sensibus corporis."

Peculiarly enough, this very tendency seemed at first to work into the hands of science. Mr. G. H. Mead has in fact said that the Christian soul was the contribution of religion to the embryonic movement of science. Partly due to Platonic influences (such as the geometrical doctrines of the *Timaeus*), partly due to advances within mathematics itself, and partly due to the lack of such instruments as the telescope and the microscope with their later enormous empirical influence, the early science of the modern world was given a predominantly mathematical and quantitative cast. The possession of a tool focuses attention upon objects upon which the tool can be used. To the young boy with a hammer the world is canalized into objects which can be hammered; to the scientist in the youth of modern science the world becomes a world of objects susceptible to mathematical treatment. There are, however, in the experienced world many qualities not susceptible to such mathematical treatment—the play of colors, the surge of emotions, the musical structure of sounds. So the richly developed Christian soul, the vehicle of mind, became the logical harbor of refuge for the homeless aspects of reality which the mathematically conceived world refused to harbor. Nature as a vast mechanico-mathematical system was closed to mind. And mind so conceived had to include the entire realm of immediate experience; all that is present, that is given, had to be regarded as mental and put in the mind. In this way the subjectivistic tendency fostered by the religious tradition of the Middle Ages became complementary to the demands of a growing science, a situation of extreme interest and of enormous significance for the comprehension of the course of Western thought.

---

[1] Cf. Windleband, *A History of Philosophy*, pp. 325, 326.

The result is clearly outlined in Galileo's contrast of the realm of primary qualities with the realm of secondary qualities: the former being the objects of knowledge, quantitative, objective, absolute, and mathematical; the latter being the realm of opinion, qualitative, subjective, relative, and sensible.[1] If Descartes really found nothing of significance in Galileo's works, as he claimed,[2] that in itself is but an indication of the fundamental nature of the view in question, for Descartes only amplifies the situation as it had appeared in his contemporary. Such great underlying philosophical structures are representative of whole cultural movements and are not merely personal idiosyncrasies of individual thinkers.

Descartes in a famous sentence says, "I was especially delighted with mathematics . . . I was astonished that foundations, so strong and solid, should have had no loftier superstructure raised on them."[3] With his conception of a "universal Mathematics"[4] Descartes hoped to raise this superstructure. He felt that all things are "mutually related in the same fashion" as the objects of geometry.[5] The external world is so conceived as to be identical with the object of geometry as interpreted by Descartes. That is, the world of matter has become a world of pure extension. Descartes had said in *Le Monde,*[1] "Give me extension and motion and I will construct the world." It is to be noted that in this account nature has become entirely mechanized. Were Descartes always consistent, he would have to say that mind can in no sense be a cause of the motions in nature. The modern world, retaining for a long period the substance view of mind, and the sharp separation of pure thought and sense, is increasingly impelled by the advance of science to give up the Greek view of mind as the principle of motion and order in nature. In this sense, the philosophy of the seventeenth century separates mind and nature more sharply than any period of classic thought.

---

[1] E. A. Burtt, *The Metaphysical Foundations of Modern Physical Science,* p. 73.

[2] Mahaffy, *Descartes,* p. 35.

[3] Haldane and Ross, *The Philosophical Works of Descartes,* Vol. I, p. 85.

[4] Ibid., p. 14.

[5] Ibid., p. 92.

With the external world reduced to pure extension, the world of sense could no longer be regarded as giving a direct contact with the external world. The external world must be known, if at all, by mind and not by sense. Even those aspects of "thought," as Descartes uses the term, which are closest to sense, such as the imagination and the emotions, must be called in question as vehicles of knowledge.[2] Paradoxically, nature is only to be known by that which is intrinsically different from nature, namely, by pure thought, by mind.

Descartes' treatment of mind is well known. From the fact that thinking occurs, on the assumption that any attribute must belong to a substance,[3] Descartes concludes that there must be a substance which thinks, and this he calls mind. Since mind and body can be conceived apart, on the assumption that things which can be thought apart are separable in reality,[1] mind and body must differ in essence. The essence of mind is thought; the essence of matter is extension. It is hardly necessary to point out today the dubious status of both assumptions. From the bare fact of thinking, nothing can be concluded. "Thinking is, therefore thinking is" is all that may be said, and not "I think therefore I am." It is clearly the religious tradition that supplies the content of the "I" as used by Descartes. This is evident when we remember that although Descartes calls mind "a thing which thinks,"[2] mind is really an aspect of the soul, in brief, it is "the whole of that soul which thinks,"[3] "a thing which doubts, understands, affirms, denies, wills, refuses, which also imagines and feels."[4] Although the soul or mind, as he here admits, imagines and perceives, these aspects of mind are somehow due to the fact that mind is "inclosed" in the material body, for the essence of mind is pure thought. Not merely is mind sharply separated from nature, but within mind itself there is an unreconciled dualism between its purely rational and its sensuous constituents. In Descartes, no more than in Plato and Aristotle, has the gap between thought and sense been bridged. In addition, there has been raised in a more acute form the relation of mind as a whole to the world. The Cartesian dualism is a double dualism.

---

[1] Mahaffy, op. cit., p. 157.

[2] Descartes, *Meditations,* VI.

[3] Haldane and Ross, op. cit., vol. II, p. 53.

There is only one more phase of the Cartesian philosophy that must be mentioned here. It is clear that Descartes' position would require a complete denial of any participation of mind in nature. Descartes, however, shrinks from this conclusion. Instead, he affirms that the mind uses the body as an instrument, and that it can act independently of the brain.[1] While there are passages where Descartes states that mind is "coextensive with the body,"[2] he continually repeats "that the soul does not perceive excepting in as far as it is in the brain."[3] The external world affects the body, the motion is transmitted to the brain, and then the mind somehow has ideas. How this takes place, Descartes admits cannot be understood.[4] In a similarly incomprehensible manner, the mind can, by acting on the pineal gland, control the motions of the body.[5] Descartes desperately tries to keep mind causally related to the world, but it is clear that this is inconsistent with his metaphysics, as Locke and Leibniz realized. The ambiguous treatment of imagination is further evidence that the interaction is never successfully accounted for.[6]

---

[1] Ibid., pp. 32, 59, 100.

[2] Ibid., vol. I, pp. 152, 153.

[3] Ibid., vol. II, p. 210.

[4] Ibid., vol. I, p. 153.

Mind remains, then, inclosed in its ivory tower, a substance different in essence from the material world which in some mysterious manner it is supposed to know. What is given in experience is merely a set of cues useful for the preservation of the organism, and not genuine aspects of the physical world.[7] When it is added that the fundamental features of this world view were adopted by Newton, it is evident that in what Mr. E. A. Burtt has well called the Galilean-Cartesian-Newtonian world-view, there had appeared a doctrine arising out of the fundamental features of Western civilization, a doctrine destined to dominate the thought of this civilization for many generations after the period of its formation. It is this view which has insinuated itself into our daily vocabulary, and into the philosophy which underlies the world-view of the science of the recent past.

---

[1] Ibid., vol. II, p. 212.

[2] Ibid., p. 255.

[3] Ibid., vol. I, p. 293. Cf. p. 196.

[4] N. K. Smith, *Studies in the Cartesian Philosophy*, pp. 82, 83.

[5] Haldane and Ross, op. cit., vol. I, pp. 289-291.

[6] N. K. Smith, op. cit., Appendix B to Chap. III.

[7] Haldane and Ross, op. cit., vol. I, pp. 194, 197.

IV

It is no exaggeration to say that the main movement of philosophy since Descartes has centered around the criticism and supplementation of the Galilean-Cartesian-Newtonian world-view. The direction of the movement has been toward the integration of mind and matter, and away from what Whitehead has called the "bifurcation of nature." As might be expected, attempts to dispense completely with the dominant world-view were preceded by a series of efforts which dealt with only single phases of the complex problem. Continuing, for the time being, to use the vague terms mind and matter, it is clear that the reactions to the dualistic metaphysics could take three directions: mind could be assimilated by matter; matter could be assimilated by mind; or mind and matter could be regarded as metaphysically parallel. Historically these possible positions are represented by Hobbes, by Berkeley and the German idealists, and by Spinoza.

Hobbes rightly insists that from the fact of thinking, Descartes could not legitimately conclude that there is an immaterial spiritual substance which thinks, thus anticipating a similar criticism by Kant. Hobbes then goes to the opposite extreme in contending that it is the physical body which thinks. All change is merely a change in the motion of bodies, and nothing can cause motion but another moving body.[1] The whole world of sense, which he calls "ideas" or "phantasms" or "apparitions,"[2] and which with speech make up all that is meant by mind,[1] exists only in the organism and not in the world of bodies. Sense "can be nothing else but motions in some of the internal parts of the sentient."[2] As Hobbes recognizes,[3] on such a view "mind will be nothing but the motions in certain parts of an organic body." Hobbes therefore keeps intact the matter side of the Galilean-Cartesian-Newtonian world-view, and its doctrine of representative perception. His innovation is in the insistence that the world of sense is itself merely a system of material motions. This world of matter is clearly taken on faith, since Hobbes never deals with the question as to how it is possible, if only the sensuous is given, to pass beyond the sensuous to a world of matter, and to affirm that the experienced world is really only material.

---

[1] Molesworth, *The English Works of Thomas Hobbes,* vol. I, pp. 69, 70, 124.

[2] Ibid., vol. IV, pp. 2, 3.

Spinoza continues to regard mind and matter as ultimate aspects of the universe, but instead of regarding them as two distinct substances, he regards them as two aspects of the single substance which can equally well be called God or Nature. The world as a whole, and every part of the world, has an aspect which is mental and an aspect which is physical.[4] *Ordo et connexio idearum idem est ac ordo et connexio rerum.* If this doctrine be taken literally, it is true that there is metaphysically no problem as to how mind is related to matter, or how knowledge is possible, but the reverse problem arises as to how error is possible. Spinoza's answer shows how deeply the Galilean-Cartesian-Newtonian world-view, and Descartes' separation between pure thought and sense experience had become embedded in the stream of thought. In substance, Spinoza answers that the human mind knows the external world only through modifications of the human body,[1] and that in knowing these modifications the external body is known,[2] to be sure, but known only inadequately, and in a confused and mutilated manner.[3] To obtain knowledge, it therefore becomes necessary to distinguish the idea as pure thought from other "modes of thinking,"[4] and to regard error as due to the imagination, that is, to the influence of the human body on the process of mind.[5] Needless to say, this solution violates the central metaphysical principle of the parallelism of mind and matter. In Spinoza, as in Descartes, pure thought remains opposed to sense, and the human mind is still separated from the external world by the human body.

---

[1] Ibid., vol. III, p. 16.

[2] Ibid., vol. I, pp. 389-391.

[3] Haldane and Ross, op. cit., vol. II, p. 65.

[4] "Thinking substance and extended substance are one and the same thing." Spinoza, *Ethics,* Pt. II, Prop. 7.

To the idealistic philosophy must go the credit for a sharp opposition to the mechanico-mathematical aspect of the Galilean-Cartesian-Newtonian world-view. It is, in fact, in this opposition that the key to the idealistic philosophy is to be found, for the historic task of idealism has been to vindicate after every great movement of science the status of human values, to resist the reduction of the complex phases of reality to the simpler phases, to restore to man a place of dignity in the cosmos. The philosophy of mind, connected so intimately with the realm of value, inevitably becomes the keystone of the idealist arch. In parrying the materialistic treatment of mind such as Hobbes would give, the traditional idealist replies with an equally extreme mentalizing of matter. Mind is the only reality; matter is but an aspect of mind or a stage in its development. In the conquest of matter idealism often has fallen back uncritically upon the theory of mind of the very world-view whose mission it has been to oppose.

---

[1] Ibid., Prop. 26.

[2] Ibid., Prop. 16.

[3] Ibid., Propositions 28, 29, 31.

[4] Ibid., Axiom III.

[5] Elwes, *The Chief Works of Benedict de Spinoza,* vol. II, pp. 31-34.

Berkeley, for instance, shows conclusively that matter cannot be conceived as a substratum of the world of experience if this be conceived as being the effect in the mind of the impact of the material world on the nerves and ultimately on the brain. The mind cannot be put in the brain, for the only brain that is known is the brain as experienced, and this brain, as a collection of "ideas" or experiences, would itself have to be regarded as mental and "in" the mind.[1] In general, Berkeley maintains that there can be no necessary inference from the given qualitative world to a quantitative world of matter intrinsically different in essence. Berkeley takes the bold step of identifying the perceived world with the "real" world of nature, and thereby becomes the patron saint of all radically empirical philosophies. There is no world more real than the world as directly given, and Berkeley is keenly aware of the reconstruction in the philosophy of science which this position entails.[2] Berkeley's own passage from phenomenalism to idealism seems to have been due to his uncritical acceptance of the then dominant view of mind as substance. Regarding all that is directly present as "in" or present to a "mind" or "spirit,"[3] mind being regarded as the unperceived perceiver which is grasped by "notions" and not by ideas,[4] Berkeley is led to the conclusion that only minds and their contents exist. The world becomes a system of minds, and any existing thing which is not an object of finite mind owes its existence to its being an object of the infinite mind or God.[5] As for the materiality of the things of daily life, such as tables and bodies and stars, Berkeley's position is naïvely realistic: these things are as they are perceived to be, even though their reality cannot be severed from the relation to mind. It is only the underlying material substratum, the philosophic and not the vulgar matter, that Berkeley's position denies.[1]

---

[1] A. C. Fraser, *Berkeley's Complete Works*, vol. I, pp. 420-422.

[2] Ibid., pp. 285, 295.

[3] Ibid., p. 470 for Berkeley's use of the phrase "in the mind."

[4] Ibid., pp. 272, 336, 444-451.

[5] Ibid., pp. 260, 261; 446, 447.

However different Hegel may be in certain respects, it is clear that there is a great kinship between his and Berkeley's view. For Hegel reality is a single process that may be equally well called Mind or Spirit or Thought,[2] and which in attaining its self-realization appears in forms which seem other than itself, but which in reality are only forms through which the ultimate self-realization is obtained.[3] All of nature, for instance, is merely the process by which the infinite mind attains its fullest concreteness.[4] Mind as a self-evolving universal has no object outside of itself to which it has to conform,—the objects which mind knows are its own products, its own self, in fact, in the form of "otherness."[5] Consciousness is ultimately self-consciousness;[6] knowledge is ultimately self-knowledge;[7] substance is ultimately self or subject.[8] In a sense the infinite mind as conceived by Hegel is the expansion to infinity of the Cartesian self,[9] a mind which now includes all other minds and all of that nature which is the object of scientific study. For Hegel as for Berkeley there is in the last analysis nothing real but mind. Speaking half jokingly, it may be said that a mind so inflated as to cover and include all reality is historically, like the ambitious child's toy balloon, ready to break. After Hegel, philosophy has, for the most part, been content to treat mind less deferentially.

---

[1] Ibid., pp. 275-279.

[2] "*The Absolute is Mind* (Spirit)—this is the supreme definition of the Absolute." (Wallace, *Hegel's Philosophy of Mind,* p. 164.)

[3] Typical passages are found on pp. 16-17, 22-23, 157, 817-818 of Baillie's translation of the *Phenomenology of Mind.*

[4] Ibid., p. 821.

[5] See Wallace's *The Logic of Hegel,* pp. 44, 47.

[6] *The Phenomenology of Mind,* p. 161.

[7] Ibid., pp. 81, 822.

[8] Ibid., p. 52.

[9] Hegel calls sensations "alterations in the substantiality of the soul" (Wallace, *Hegel's Philosophy of Mind,* p. 178), and speaks of the ego as "one and uncompounded" (ibid., p. 179).

Before turning from Hegel, there is one point that may be noticed. At first reading, Hegel's works seem to be saturated with an atmosphere of movement and development. Hegel, in fact, states[1] that "mind is essentially active . . . It is wrong therefore to take the mind for a processless *ens*. . . . The mind, of all things, must be looked at in its concrete actuality, in its energy." This view of mind as process becomes increasingly prominent in the later idealistic developments, but in Hegel it is checked by the doctrine that all development is the development of a universal or identity that persists through the whole process. Although the absolute mind is "always identity returning and ever returned into itself," it is also a "self-centered identity."[2] It is "pulsating within itself, but ever motionless, shaken to its depths, but still at rest."[3] Further, the development is not a development in time,[4] so from one point of view mind is not a process at all but a completely realized non-temporal Absolute which, as Emerson would say, "lies sheltered in perfect repose." So that although Hegel suggests the view of mind as process, he himself never completely deserts the substance conception of mind, and to that degree never completely frees himself from one feature of the world-view which he so earnestly opposed.

[1] Wallace, *The Logic of Hegel,* p. 69.

[2] Wallace, *Hegel's Philosophy of Mind,* p. 291.

[3] *The Phenomenology of Mind,* p. 157.

[4] Ibid., p. 813.

In the above historical sketch, no reference has been made to Hume or to Kant. Although both of these thinkers remain within the framework of the Galilean-Cartesian-Newtonian world-view, both suggest, without development, an approach to the problem of mind that is so much in the spirit of later views that it has seemed advisable to sacrifice historical to logical considerations.

Hume begins his account with the current doctrine that given experiences are "perceptions of the human mind."[1] He continues to think of experience as caused by material objects working through the senses to produce impressions in the mind,[2] maintaining specifically that there is no reason why matter could not cause thought.[3] But Hume's analytical acumen forces him to realize that such material objects cannot be philosophically justified, and that there is no evidence whatever for a substance view of mind. He accordingly rejects the view of mind as substance and is even led to say that "nothing appears requisite to support the existence of a perception."[4] Hume's scepticism and his turn from metaphysical to social and moral considerations are but the expression of the fundamental conflict in Hume between the Galilean-Cartesian-Newtonian world-view, and the empirical attitude, a conflict which Hume never resolves. It is while in the empirical mood that Hume suggests another distinct approach to mind, namely, the view of mind as a relation of things that in themselves are not necessarily mental. Instead of regarding experience as mental or in a mind, he suggests that mind is a relation or grouping of experienced factors which may pass in and out of mind unchanged. "A mind,[5] is nothing but a heap or collection of different perceptions, united together by certain relations, and supposed, though falsely, to be endowed with a perfect simplicity and identity." "They are the successive perceptions only, that constitute the mind."[1] "The mind is a kind of theatre . . . There is properly no *simplicity* in it at one time, nor *identity* in different." Hume never reworked his philosophy from the basis of a relational theory of mind,[2] in fact he rejected the view because he regarded perceptions as dependent upon our organs,[3] a rejection that is entirely beside the point, and which merely shows how deeply seated in Hume was the doctrine of a material substratum working upon the body to produce "mental" effects. While passing beyond the substance view of mind, Hume never adequately explored the implications of the alternative relational view which he suggests but dismisses.

---

[1] *A Treatise of Human Nature,* Bk. I, Pt. 1, sec. 1.

[2] Ibid., Bk. I, Pt. II, sec. 4. Selby-Bigge's edition, p. 64.

[3] Ibid., Bk. I, Pt. IV, sec. 5; p. 250.

[4] Ibid., p. 234.

[5] Ibid., sec. 2, p. 207.

A somewhat similar situation is revealed in Kant. Although Kant continues to believe personally in the kind of a soul or self that can be immortal, and although there can be but little doubt that it was this transcendent soul which was the seat of the forms of perception, the categories, the ideas, and the unity of consciousness, which was, in short, the synthetic mental activity which Kant introduced to justify certain and yet objectively relevant knowledge,[4] Kant's own analysis of the so-called rational psychology, revealing, as Hume had done, the paralogism of the Cartesian passage from the fact of thinking to the existence of a unitary spiritual thinker, forces him to admit that the substance view of the soul and mind cannot gain entrance into a critical philosophy. The way is accordingly open for an empirical analysis of the synthetic function of mind—a way which Kant himself did not follow, but which becomes of increasing interest after the notion of mental synthesis had received its romantic metaphysical elaboration by the romantic idealists. In defending himself from the charge of subjective idealism, Kant goes so far as to admit that consciousness of self itself requires a consciousness of external objects.[1] The development of this suggestion would dispense with a mind lying in wait to impress unity upon a world that in itself is merely a cosmic kaleidoscope, and would lead to a genuinely Copernican revolution instead of the Ptolemaic return which Kant mistook for a Copernican advance. In Kant as in Hume the movement to a critical empiricism is hindered by the retention, as over-beliefs to be sure, of certain features of the Galilean-Cartesian-Newtonian world-view. In Hume it is the ghost of matter that lingers; in Kant it is the ghost of the mind.

---

[1] Ibid., sec. 6, p. 253.

[2] Berkeley, in the *Commonplace Book,* had suggested that "Mind is a congeries of perceptions." (Fraser, op. cit., vol. I, p. 27) Woodbridge remarks that the term perception was used so vaguely that it is doubtful whether mental acts or experienced content are referred to in this sentence (*Studies in the History of Ideas,* vol. II, p. 207, note). Due to other remarks of Berkeley (see Fraser, op. cit., pp. 33, 34) it seems reasonable to suppose that Berkeley anticipated Hume, but like Hume, never carries out the suggested implications.

[3] See the *Treatise,* Bk. I, Pt. IV, sec. 2, pp. 210-211.

[4] See N. K. Smith, *A Commentary to Kant's Critique of Pure Reason,* pp. 261, 476. Kant speaks of the "unity of consciousness which precedes all data of intuition, and without reference to which no representation of objects is

In closing this brief and too condensed historical survey, let us merely recall that in the early period of speculation in which mind and nature had not been sharply separated it was possible to regard mind as the principle of order and

---

possible." (Müller, *Kant's Critique of Pure Reason*, p. 88; *cf.* p. 694) In the *Dissertation* he had stated that "all sensitive apprehension depends upon the special nature of the subject, in so far as it is capable of being modified in diverse ways by the presence of objects." (J. Handyside, *Kant's Inaugural Dissertation*, p. 44).

[1] Müller, op. cit., pp. 779-781. See Prichard's comments, pp. 321 ff., *Kant's Theory of Knowledge*. For a good discussion of the central conflict in Kant see Chap. VII of N. K. Smith's *Studies in the Cartesian Philosophy*.

motion in nature; that as the difference between mind and matter developed into the opposition of two substances, mind was withdrawn from nature, and the gap between pure thought and matter became absolute; that as the attempts to reduce mind to matter, or matter to mind, or to conceive mind and matter parallelistically were found to introduce new difficulties, a tendency appeared to shift from the category of substance to the categories of process and relation, and to shift from metaphysical to empirical treatments of mind. In the contemporary period this shift is, with certain exceptions, accelerated, and it now becomes advisable to subject the substance, process, and relation theories of mind as they appear in contemporary thinkers to a more detailed analysis.[1]

---

[1] Space did not permit a discussion of the views of Locke and Leibniz. Locke represents the mature development of the Cartesian world-view, except that the empirical emphasis bridges the gap between pure thought and sense that is found in Descartes. The dualism between the world of experience and the world of matter remains, and the difficulties of this dualism begin to make themselves felt. Leibniz, in denying that perception could be explained by matter (*Monadology,* Sec. 17), helped to transfer the discussion to experiential terms, and the doctrine of the self-development of the monads helped to determine the emphasis of later German philosophy upon the creative activities of the subject, setting the pattern for the interpretation of the material world as a stage in the psychical development of the subject.

II

THREE THEORIES OF MIND

I. MIND AS PROCESS

IN discussing Hegel, it was observed that one phase of Hegelian philosophy leads into the view that mind is a self-evolving process. This doctrine of mind as a process whose content is the world has become more and more dominant in the idealist philosophy and may, in fact, be taken as the central doctrine of the idealist philosophy of mind. The absolutistic element that is prominent in Hegel, that is, the refusal to take time seriously by considering the development of mind to be logical and not temporal, has gradually given way in all countries to a more genuinely activistic interpretation of the mental life, which has not, however, sacrificed the emphasis upon the autonomous character of mind. "Idealism," Gentile writes,[1] "is the rejection of any reality which can be opposed to thought as independent of it and as the presupposition of it." To exhibit the idealist theory of mind more completely, it will be well to glance at the earlier post-Hegelian type of theory as it is revealed in Bosanquet, and the most recent form of the theory as it appears in Gentile.

Bosanquet, like Hegel, reveals a passion for concreteness and continuity, and a dislike for all dualistic modes of interpretation. Influenced by admiration for the integral wholeness of the life of the Greeks, there burns beneath his words a constant love for the whole of things. This sense of the whole was clarified for him in the experiences of art, and morality, and religion, which become foretastes, as it were, of the systematic absolute whole that reveals itself through finite centers, but which in itself is a realized total perfection. The finite centers can only approach this perfection of wholeness but cannot attain it in its completeness.

---

[1] *The Theory of Mind as Pure Act,* p. 18. General statements of the idealist position may be found in G. W. Cunningham's *Five Lectures on the Problem of Mind,* and H. W. Carr's article on "Idealism as a Principle in Science and Philosophy," found in Muirhead's *Contemporary British Philosophy,* vol. I.

The same duality of emphasis upon realized perfection and temporal development appears in Bosanquet that appeared in Hegel. Although for Bosanquet change, like morality and religion, is ultimately only finite appearance, since "the whole cannot change,"[1] nevertheless, in the desire to keep the reality of moral struggle and to do full justice to the finite, Bosanquet insists upon change within the whole. "The whole," he writes,[2] "is such as on the one hand to include change, and on the other hand not to break away from totality." While it cannot be said that Bosanquet succeeds in reconciling perfection and development, it is in the tendency toward the development of greater wholeness that Bosanquet's theory of mind is to be found.

On this theory, "mind has nothing of its own but the active form of totality; everything positive it draws from nature."[3] He therefore insists upon the "continuity of the real world with mind:"[4] "thought moves in the world of real objects, and has never come out of it."[5] "There is nothing to be called 'the mind' of which there are empty acts exercised upon objects,"[6] for "the mind is a number of things growing, interworking, progressing," "On all sides," he writes,[1] "there is a convergence on the conception of mind as consisting of what it does and experiences." "Thought which deals with no given, and constructs no order is a *res nihili*."[2]

---

[1] *The Meeting of Extremes in Contemporary Philosophy,* pp. 176, 177.

[2] Muirhead, op. cit., p. 73. This whole essay on Life and Philosophy gives an excellent approach to Bosanquet's thought.

[3] *The Principle of Individuality and Value,* p. 367.

[4] *The Distinction Between Mind and Its Objects,* p. 49.

[5] *Three Chapters on the Nature of Mind,* p. 113.

[6] Ibid., p. 20.

While insisting that mind cannot be divorced from the world, Bosanquet insists equally strongly that the world cannot be divorced from mind. In good Hegelian fashion he argues[3] that "all that is an object of thought is ultimately a constituent of it," that "thought is a total and continuous judgment which sustains by affirmation the world of reality," that even an experience of blue "must have in it the life of mind,"[4] that "there can be no concrete whole but a whole centering in mind,"[5] no world without "consciousness as its centre."[6]

It is evident, I believe, that Bosanquet's theory of mind takes two directions. From one point of view, mind practically becomes the active aspect of experience, the tendency to the formation of more and more comprehensive wholes. It is while thinking in this direction that Bosanquet talks of the "energy of thought" as "the nisus to the whole."[7] But as the above quotations show, Bosanquet not only makes mind the active aspect of totality, but explicitly identifies mind with "the conditions of totality."[8] Mind is then not merely the active aspect of reality but is somehow the very condition of reality. On the first type of treatment, the

---

[1] *The Meeting of Extremes,* etc., p. 25.

[2] Muirhead, op. cit., p. 60.

[3] *The Meeting of Extremes,* pp. 23, 22.

[4] *The Distinction Between Mind and Its Objects,* p. 33.

[5] Ibid., pp. 39, 40.

[6] Ibid., p. 58.

[7] *Three Chapters on the Nature of Mind,* p. 70.

[8] *The Distinction Between Mind and Its Objects,* p. 44. Hoernlé does not discuss this aspect of Bosanquet's theory in his treatment of Bosanquet in *Idealism as a Philosophy,* pp. 236-243, 264-267.

Absolute, as the whole, could have no nisus to wholeness and could not be regarded as mind; on the second type of treatment, if every whole must center in mind, the absolute whole may be regarded as an absolute mind. In one case mind is an aspect of reality; in the other it is reality. It is undoubtedly possible to reconcile these two aspects of Bosanquet's treatment. They have been separated only in order to focus attention upon the fact that the view of mind as process need not be carried to the position of the idealist that the world ultimately is mind.

With Bosanquet's refusal to separate mind from the world few contemporary philosophers would quarrel, but many would hesitate to agree that the world needs mind as its center and as its condition of self-existence. What argument can the idealist use to support his contention that the world is mind-dependent? Bosanquet makes very little use of the older Berkeleyan type of argument that nothing can be experienced or conceived without thereby relating it to a subject. Such considerations show merely that nothing can be experienced without being experienced, and do not show that all reality must be experienced to exist, or that the experienced object is mental.[1]

Bosanquet's argument, in so far as it can be called an argument, seems to be based upon the relativity of observed content to the observer. To the "open-door" policy of the new realism Bosanquet replies,[2] "your organism, which you cannot separate from its mind, is one of the conditions which things require for the manifestation of their complete being."

---

[1] Perry's discussion of this argument under the name of the "ego-centric predicament" has served to focus the question, even though the logical fallacy of accident is all that is involved. See discussion on pp. 129-132 of his *Present Philosophical Tendencies.* Hoernlé, op. cit., p. 114, seems to accept the brunt of Perry's criticism.

[2] *The Distinction Between Mind and Its Objects,* p. 48.

In this retort, it appears that Bosanquet is regarding relativity to the observer as a case of relativity to mind. This same type of argument is employed by certain idealists who attempt to use the theory of relativity as support for an idealistic metaphysics. This argument, however, has no force unless one is already committed to an idealist theory of mind, for it does not necessarily follow that reality is mind dependent even though the experienced world be dependent upon the organism. A holder of a relational theory of mind could regard experience as dependent upon the organism, but regard mind as dependent upon these experiences rather than the experiences dependent upon mind. Similarly, the pragmatist is often inclined to regard experience as dependent upon the organism without feeling any compulsion to subscribe to the idealist theory of mind. Not merely does the idealist argue invalidly from the ego-centric predicament, but is also liable to confuse the relativity of experienced content to the organism with the dependence of this content upon a mind. A consideration of the relativity involved in the operation of a camera should help to show the possibility of other interpretations.

The main approach of the idealist to the problem of mind is, however, metaphysical rather than analytical. Before glancing at the problems which this metaphysics proposes, it may be well to look at its development in the contemporary Italian philosopher, Gentile, who has even accentuated the reaction against the static aspects of idealism which is found in Croce. Gentile's idealism regards reality as a concrete universal mind which generates all objects from its own activity. Thinking is the whole or absolute reality, "provided we mean transcendental and not empirical thought."[1] Gentile's idealism is essentially dynamic and opposed to the slightest remnant of even a mental substance. Mind is "pure act" and "has no existence apart from its manifestations."[1] His view "resolves the world into spiritual act or act of thought."[2] "In so far as the subject is constituted a subject by its own act it constitutes the object."[3] As applied to the individual this means "I am only in so far as I think."[4] That is, the mind does not exist as a substance before its act, but the mind as subject becomes a reality at the same moment that it constitutes its object a reality. In summarizing his position Gentile writes[5] that "mind, the spiritual reality, is the act which posits its object in a multiplicity of objects, reconciling their multiplicity and objectivity in its own unity as subject. It is a theory which withdraws from mind every limit of space and time and every external condition."

---

[1] *The Theory of Mind as Pure Act*, p. 5.

In Gentile's position the theory of mind as process has perhaps reached its most extreme position, so we may merge our criticism of his position with the discussion of the process theory in general. A universal developing its identity through change presents serious enough difficulties, but a conception of mind as a pure act producing itself and the world in an eternally self-creative process, free as it is from the last traces of a substantive metaphysics, seems to border on the incomprehensible. A pure act, an act that is not the act of something but rather the act by which an actor appears, seems to be almost a pure nothing.

It is true that there is much loose talk in popular discussions of science that seems at first sight to substantiate such an unsubstantial position: we are told that the atom is only what it does, that an electron would become nothing if it could be stopped, that the substance of the atom therefore depends on its motion, that matter has been reduced to pure force or pure energy. But a critical science must recognize that such statements are merely vivid forms of speech. No critical physics can talk about pure force, for force is itself defined through the change of the condition of a physical body. A body that increases its velocity or acceleration or direction of motion is said to be acted upon by a force, but the force is precisely the change of velocity, acceleration, or direction of motion and not a pure force that produces the change and creates the body. Force and energy are ultimately adjectives descriptive of certain aspects of the experienced world, and not nouns representing either ultimate substances or pure acts. And as for the claim that the very being of an electron is dependent upon its motion, it must be remembered that in the theory of relativity any particular body can equally well be regarded as at rest or in motion. So if science wishes to keep both propositions it must say that atoms considered in motion constitute the substance of the world, but since any particular atom may be considered at rest, the substance of any particular body can also be regarded as nothing. This conclusion may interest the speculative thinker who wishes to avoid the very real problems that the atomic theory raises, but the proposition that the ultimate somethings are equally well ultimate nothings can hardly be said to represent a tenable scientific position. Questionable metaphysics is not to be accepted merely because its origin is labeled "science" instead of "philosophy."

---

[1] Ibid., p. 20.

[2] Ibid., p. 265.

[3] Ibid., p. 18.

[4] Ibid., p. 100.

[5] Ibid., p. 241.

Similar considerations apply to pure activity. While we may legitimately talk descriptively of concrete situations manifesting activity, we cannot regard the situation as produced by the activity. Although a subject cannot be regarded as a substance separate from its activity, the activity cannot be regarded as a pure activity which constitutes the subject. Substance cannot be separated from activity nor reduced to activity. The only mind that is known empirically is active but not pure activity, and what is true of empirical thought must be true of transcendental thought if the use of thought in this latter connection is to have any meaning at all.[1]

The doctrine that the world is ultimately mind or spirit (and the idealist does not often separate the two categories), due to its historical associations, seems to remove the harsh foreign aspects of reality, to reduce the world to material for the satisfaction of man's own longings and aspirations, to awaken, in short, a combination of humanistic and religious emotions. Whatever may be said of this emotive appeal, it is very doubtful if the proposition "the world is nothing but mind," or the proposition "the world is nothing but matter," has any genuine philosophical significance, for in either case the same further distinctions must be made: on the one hand it is necessary to recognize the distinction between a stone and the thought processes of man, only the distinction is called a difference of levels within mind; on the other hand it is also necessary to recognize the distinction between a stone and the thought processes of man, only now the distinction is called a difference of levels of matter. If it be maintained that the difference in attitude between mentalism (which has been but need not be the position of idealism) and materialism is the difference between assimilating the lower to the higher or the higher to the lower, it may be doubted if either assimilation is possible, since both sides must admit that the higher manifestations are inseparable from the lower and yet must be distinguished from the lower. With the appearance of the doctrine of emergent evolution, the old controversy between materialism and mentalism bids fair to lose both its significance and its existence.

---

[1] A defense of the position that "thinking, the pure act, is original and the thing thought is derived" may be found in Carr's article, p. 116, of vol. I of Muirhead's *Contemporary British Philosophy.*

In summary, idealism, with its theory of mind as process, as systematic development, has forced philosophy to recognize that mind cannot be separated from the world of nature, that mind is more than a mere collection of separate units or a passive receptacle, and finally, that mind cannot be reduced to the other phases of reality. It may be agreed that no theory of mind which does not profit by these results will prove to be adequate. However, in avoiding the dualism between thought and sense, mind and nature, experience and reality, idealism over-generalizes its case in regarding mind as the sole reality. Not merely is this position open to the criticisms that have already been suggested, but such a wide use of the term mind throws no light on the specific problems that were suggested as valid criteria of a theory of mind. It gives no insight into the phenomena of language, no appreciation of the mathematical type of thought, no concrete discussion of the empirical reflective process, nothing but the general hint of development in approaching the anthropological and evolutionary phases of mind. In its intoxication with the emotional reverberations of the term mind, idealism is in danger of remaining content with general propositions concerning mind. Its very passion for metaphysical concreteness may reveal itself in an inverse contentment with empirical emptiness.

II. MIND AS RELATION

It was inevitable that the idealistic attempt to interpret reality in terms of mind should meet with strong opposition from those thinkers who voiced the aspirations of the modern scientific attitude. It was out of this opposition that the realistic and pragmatic movements of philosophy emerged, the realists being on the whole influenced by the physical and mathematical sciences, the pragmatists by the biological, psychological, and social sciences. With historical relations to Hume, and influenced by Brentano and Meinong, there grew up in England about the beginning of the present century the movement known as the new realism, well exemplified in G. E. Moore. In America, with lines of influence from Mach and Renouvier, there grew up around Peirce and James a reaction that developed both in the direction of new realism and in the direction of pragmatism. For the time being we will consider only the new realistic position.

Under the general doctrine of the externality of relations, the new realism claimed that things pass in and out of relations unchanged, and that they therefore passed in and out of the relation to mind without change. Because things can only be known by being brought into a cognitive relation does not prove that only known things exist; because human thought can only deal with experienced things does not mean that these things as experienced are not independent of their being experienced. The new realism therefore passes to the position that what is known is non-mental and is independent of its relation to mind. What then is the mind?

Two tendencies in the interpretation of mind appeared in the movement: first, the attempt to regard mind as an "act" of awareness whose content is non-mental. This separation of mental act from non-mental content is characteristic of G. E. Moore and the British new realists, and will be discussed in another connection; and second, the attempt by the American new realists and later by Bertrand Russell to conceive of mind as a relationship between things not in themselves mental. It is this relational approach to mind that must now be considered.

A typical position is found in Holt, who regards mind as a grouping of objective entities, not a substance in which thing are, nor a process in which things become what they are. Even illusions are interpreted as physical and objective. Consciousness "is *not* in the skull, but is 'out there' precisely wherever it appears to be."[1] Mind or consciousness "is a cross-section of the universe, selected by the nervous system. The elements or parts of the universe selected, and thus included in the class mind, are all elements or parts to which the nervous system makes a *specific response*." "Consciousness is, then, out there wherever the things specifically responded to are."[2] Perry's position is influenced by, and very similar to, Holt's position, in that he regards mind, and also matter, as groupings of neutral entities. Since the same contents may appear in various groupings, minds may overlap, so that instead of being subjective, "mind is open to mind."[3]

Other of the American new realists take somewhat different approaches. Montague holds[4] that "all matter is instinct with something of the cognitive function . . . every objective event has that self-transcending implication of other events which when it occurs on the scale that it does in our brain processes we call consciousness." "Consciousness is the potential or implicative presence of a thing at a space or time in which that thing is not actually present."[5] Pitkin, using an analogy from projective geometry,[6] regards the field of consciousness as a projection of the environment (the projected complex) upon the reacting organism (the projectorial referent). But whether regarding mind as a grouping of neutral entities, or an implicative presence, or a projection field, the new realists unite in agreeing that mentality is not a characteristic of things in themselves, but of things only in relation to the reacting organism. There is no mind apart from things in relation.

---

[1] *The New Realism,* p. 353.

[2] Ibid., p. 354.

[3] *Present Philosophical Tendencies,* Chap. 12.

[4] *The New Realism,* p. 283.

[5] Ibid., p. 281.

[6] Ibid., pp. 443 ff.

On the relational view of mind, as on the idealistic view, the content of mind is the content of the world. Unlike the idealist, however, the new realist regards no content of mind as dependent upon mind for its reality. Since not even a mental act is admitted by the American new realists, logic and mathematics cannot have anything mental about them. Mathematical propositions must be part of the content of the world which can enter into the mental relation; they must be discovered rather than created. Marvin tells us[1] that the proposition "$2+2=5$" was false fifty million years ago. The new realism has been forced to adopt a Platonic realm[2]—the realm of subsistence—to house the contents of logic and mathematics. In a similar manner, error, hallucinations, and illusions must be regarded as non-mental, as objects rather than products of mind.

We may pause a moment to envisage the difficulties in regard to error which this view encounters. Holt regards error as simply the beholding of objective error, truth the beholding of objective truth, where objective error is regarded as objective contradiction.[3] In spite of Holt's specific consideration of the point at issue, few thinkers have been able to convince themselves that the objective world can rightly be said to exhibit "contradiction," a term which seems to apply more to the realm of discourse than to the relations between things. To explain error, Perry, Montague, and Pitkin frankly fall back upon some form of dualism: Perry[1] insisting upon the possibility of selecting neutral events to produce a "subjective manifold" that may differ from and be mistaken for the groupings which make up the physical world; Montague[2] regarding error as due to the uncorrected "distortion" somewhere introduced into the physical process of being affected by external things; Pitkin[3] interpreting error as being due to entities that are "projectively indiscernible," that is, to regarding a present content as being projected by an object which is not really "projecting." While all of these views present difficulties that cannot be considered in the present connection, it may be noted that the attempt to separate error from mental processes is fraught with difficulty, and that the introduction of some dualism within the world makes it difficult to maintain that the world as present is the world as it is, for if things always are as they appear to be, error is impossible. In attempting to explain error the new realist is in danger of giving up his cardinal principle. Pitkin frankly says[4] that the view that "consciousness is the mere knowing of physical things as they are 'in themselves' " is a "naïve doctrine." On the basis of such statements, it is safe to say that within the new realism itself appeared the seeds out of which the temporally later movement, critical realism, was to grow.

---

[1] Ibid., p. 53.

[2] Ibid., p. 35.

[3] Ibid., p. 357 ff.

The most thoroughgoing relational view of mind is found in the new realistic stage of Bertrand Russell's philosophical development. Renouvier, whom Höffding has called "the patriarch of contemporary philosophy," passed through a phenomenalistic period[1] in which he banished any substance underlying the given world, thus making the phenomenal world the very stuff of reality, antedating all distinctions between subject and object, perceiving and perceived. Similarly, Mach based his doctrine upon a pure phenomenalism, claiming that the data of physics and psychology are the same in content, and differ only in the relations of this given content. The physical and the mental are determined by different systems of elements, not by different elements. "A color is a physical object as soon as we consider its dependence, for instance, upon its luminous source, upon other colors, upon temperatures, upon spaces, and so forth. When we consider, however, its dependence upon the retina . . . it is a psychological object, a sensation."[2] Sensations are not in the head, rather the " 'head' shares with them the same spatial field."[3] "There is no rift between the psychical and the physical, no inside and outside, no 'sensation' to which an external 'thing' different from sensation, corresponds." "The world of sense belongs both to the physical and the psychical domain alike."[4] Mach has thus substituted a thoroughgoing phenomenalism or radical empiricism for the dualism of the Galilean-Cartesian-Newtonian world-view.

---

[1] *Present Philosophical Tendencies,* pp. 324, 325.

[2] *The New Realism,* pp. 286-300.

[3] Ibid., p. 459.

[4] Ibid., p. 463.

Mach has been mentioned because although Russell's position is not phenomenalistic,[5] it shares with Mach the doctrine that mind and matter are but different relations between elements that in themselves are not mental nor material, and the doctrine that psychology and physics do not deal with objects that differ in substance.

---

[1] Renouvier's later philosophy is idealistic. See Gunn, *Modern French Philosophy*, p. 116.

[2] *The Analysis of Sensations*, Open Court edition, 1914, p. 17.

[3] Ibid., p. 27.

[4] Ibid., p. 310. The definition of mind in the existential psychology is given in words that could have been written by Mach. Thus Titchener defines mind as "the sum total of human experience considered as dependent upon the nervous system." (*A Text Book of Psychology*, p. 16).

[5] Russell, *The Analysis of Matter*, p. 399.

For Russell the ultimate aspects of reality are what he calls "events," which are happenings finite in all dimensions. The clearest examples of events can be taken from the realm of percepts—sounds, flashes of lightning, and the like. On various grounds Russell is led to postulate other events besides percepts. Events are so related that one event is connected with other events in space and time, and some events may be regarded as "appearances" of other events. Thus there are any number of unperceived events that may be regarded as "appearances" of a star.[1] If we collect the appearances of different stars at a given place where there is a brain, such a collection is a mind (a percept, which Russell, unlike Mach, always calls mental,[2] is "the appearance of the object from a place where there is a brain"),[3] while if all the appearances of a particular star are collected at different places, such a collection is a physical object, in this case, the star itself. A mind so defined is simply a collection of sensations which are mental only in virtue of their relation to the organic brain.[4] There is nothing in mind but the sensations so related. Just as there is no "the" real table, the table being simply the whole set of correlated actual and possible sensations,[5] so there is no mind substance as contrasted with the set of sensations.

It would be unwise to follow further Russell's account, which becomes complicated by an increasing reference to psychological and biological matters and which soon deserts, as will be seen later, the new realistic position. Instead it is necessary to bring the discussion of the relational theories of mind to some sort of conclusion.

---

[1] See *The Analysis of Mind,* pp. 99-107.

[2] *Philosophy,* p. 286.

[3] *The Analysis of Mind,* p. 131.

[4] Ibid., p. 69.

[5] Ibid., p. 98.

The most striking characteristic of the relational theories of mind is their sharp contrast to the idealistic formulations. While the idealist expands mind to include all reality, the new realist so contracts mind that there is nothing mind can really call its own, not even itself. Reduced to either a bare awareness or a group of sensations in a certain relation to the brain or nervous system, mind has become in the hands of the new realists a very feeble and forlorn aspect of reality. All the contents of the Cartesian mind have been returned to the world, and mind has become merely a relation between the very contents it had previously so proudly claimed. Certainly if idealism had unduly magnified mind, the new realism has unduly minimized it.

It is to be noted that this type of theory makes practically no reference to those aspects of experience which would normally be called the higher mental processes, that is, the judgments, and the chain of thoughts usually evoked in any attempt to deal with truth and error. The very kind of "mental" activity necessary to produce such a theory of mind is neglected in the theory itself, a neglect that is ultimately due to the neglect of psychological considerations by the new realistic movement. With mind reduced to such a shadow, all such phrases as "he has a better mind than x," "y is capable of hard mental work," "the evolution of the animal mind," can only apply to the selectivity of the nervous system singling out neutral elements! I am afraid that we must apply Broad's term here—this is a "silly" theory. With mind reduced to a bloodless shade, the things squeezed out of mind must be returned to the world, and those which the realm of matter will not accept, must be enthroned in the realm of subsistence. Here all truths and errors, all logical and mathematical systems and propositions lie in wait for the nervous system which can bring them into the light of experience. Fortunate for Newton that his nervous system selected out the timeless truth of the law of gravitation, and the propositions of the calculus; fortunate for Cantor that his nervous system selected out the transfinite cardinals and brought them into human vision! The history of the human mind becomes the history of human nervous systems.

As for Russell, in so far as his doctrine is unambiguous, it must be said that *The Analysis of Mind* is almost a treatise on how to analyze mind away. The view that mind is a collection of sensations dependent on a living brain gives no insight into the systematic activity of mind which idealism has rightly emphasized. What of the gigantic mental labor that Russell performed in the elaboration of *Principia Mathematica?* Was that merely a collection of sensations at a place where there was a brain? What of any proposition in *Principia?* Does not its very nature and meaning depend upon its being taken as a whole, that is, as a proposition, and not broken up into a collection of sensations? Although Russell has given less and less importance to the realm of subsistence which enveloped his earlier mathematical philosophy, it is difficult to see that Russell has provided any place in his later[1] views of mind for the type of mental activity that a genuine mathematician performs.

New realism did well to recognize that mentality is not a characteristic of experience content *per se,* but it failed to find a satisfactory criterion for the appearance of mind in experience. New realism did well in attempting to state as much of reality as possible in non-mental terms, but it appears to have carried the attempt farther than analysis justifies. If to the idealist we say "more matter, never mind so much about mind!", to the realist we must say "more mind, never matter so much about matter!" And history appears to have sustained this verdict.

III. MIND AS SUBSTANCE

It is very difficult to define adequately the term "substance," and an adequate definition will not be attempted. Roughly, it may be said with Broad[1] that "an existent is more of a substance the longer it lasts and the less dependent it is on anything else," that is, "endurance and capacity for independent existence" may be taken "as two tests for substantiality."[2] The concept of substance is undoubtedly derived from the semi-permanent things of daily life, tables, rocks, organisms, and the like, all of which are called material substances because they can be touched, handled, weighed. By an intellectual refinement it is possible to conceive of material substances which would endure throughout time and which would persist unchanged even if the remainder of the universe were destroyed. By a further refinement it is possible to conceive of substances which would have endurance and independence without having the differentiating attribute of materiality, and which could not therefore be touched, weighed, or handled. Mind and soul were previously conceived of as such immaterial substances.

The theory of mind as an immaterial substance need not detain us long. Not merely do few philosophers of importance now hold the doctrine, but two weighty reasons make its retention difficult though not impossible. In the first place, both Hobbes and Kant rejected as invalid the transition from the fact of thinking to the existence of an immaterial substantial mind that thinks, and their reasoning has never been successfully challenged. In the second place, there is no empirical evidence for such a mind. Minds do have a certain endurance in time just as bodies do, but the evidence is all against a mind that is independent of the other factors of experience. As empirically known, mind is inextricably linked with the whole of the appearing world. No immaterial mind, independent of the rest of the experienced world, is found, and empirically there is nothing more to be said. Of course there *may* be such a mind, and the moon *may* be made of green cheese.

---

[1] *The Mind and Its Place in Nature,* p. 31.

[2] Ibid., p. 33.

In contemporary thinkers, the attempt to conceive of mind as a material substance, or attribute of such a substance, is more common. With the prevalent biological cast of modern thought, it is natural that the human body or some portion of the body, such as the brain, should be identified with mind.

In this connection a consideration of the views of Samuel Alexander is instructive. At the outset it must be admitted that Alexander's account brings together several streams of thought, and can only in part be regarded as representative of a substance theory. Unlike most new realists, Alexander rejects the relational view of mind as it is found in Holt, claiming that it does not do justice to the fact that *we* are aware of what is before *us*.[1] He therefore falls back upon Moore's distinction between mental act and non-mental content.[2] In any experience, it is claimed, there are two distinct elements, the object of awareness and the awareness of the object. The first of these Alexander calls "contemplation," the second "enjoyment." "The mind enjoys itself and contemplates its object. The act of mind is an enjoyment; the object is contemplated."[3] These two aspects of experience are related by being together or "compresent." Mind cannot be an object to itself, cannot contemplate itself:[1] so-called introspection does not make the mind an object to itself but merely clarifies the enjoyment.

---

[1] *Space, Time, and Deity,* vol. II, pp. 109-115.

[2] G. E. Moore, *Philosophical Studies,* pp. 20, 24 ff.

[3] *Space, Time, and Deity,* vol. I, p. 12.

We may perhaps pause a moment at this place to consider the doctrine of mind as act aware of content. It must be admitted with Alexander that it is possible for us in looking at a horse to experience the fact of our togetherness with the horse,[2] but the question arises as to whether this togetherness is the compresence of a mental act with non-mental content, or whether it is merely the compresence of two contents. We are told to[3] "seek for the enjoyment as something which you mind or live through, and which you are . . . and . . . you will assure yourself of the compresence of the non-mental object with your enjoyed mind." But many who have earnestly sought have not been able to assure themselves of the mental act. We have seen that the American new realists have never accepted the doctrine of mind as act, and Russell, greatly influenced by G. E. Moore in the early stages of his thinking, has come to reject the doctrine.[4] Many people in searching for the bare act of awareness seem to have had the experience of Hume and James in searching for the self, the experience of finding only more content. The mental act must be found in experience, but since it cannot be found as an object, that is, cannot be contemplated, it is difficult to know whether it is found or not.[5]

We should note, however, the remarkable change in perspective that this distinction between mental act and non-mental content affords. The entire apprehended world immediately becomes non-mental and independent of mind, and it is this which distinguishes the realist doctrine of the act from the idealistic doctrine. Whatever is seen is not a part of mind, only the seeing of it is mental. There is no problem as to how we can get out of mind, for we are already "out" in the simplest experience. Even the so-called mental image is no longer mental, for only the awareness of an image is mental. As contrasted with the subjectivity of experience in the Galilean-Cartesian-Newtonian world-view, on the present position the entire content of experience is returned to the world, mind and consciousness remaining merely as the apprehension of this content. For Alexander, even memory and anticipation involve merely the compresence of mind with objects that are removed in time and space.[1] Alexander states[2] that "in the memory-image of my friend I have before my mind the revelation of my friend just as much as I have a revelation of him when I see him." The mind, then, can be compresent with other aspects of reality beyond those present in sense data. Knowledge, in fact, is merely the togetherness of mind with some factor of reality. "Whenever a mental process exists in compresence with some existent of a lower order, it is aware of that existent which is its object."[3] "To know anything is to be along with it in Space-Time."[4] The simplification of such a theory is so great, that one is tempted to embrace the theory even if the evidence for the mental act is questionable. Why furnish another example of what is said to have been Herbert Spencer's conception of a tragedy, namely, a theory ruined by a fact? The appearance of other difficulties in the theory, however, do not make so heroic a course advisable.

---

[1] Ibid., p. 17.

[2] Ibid., p. 21.

[3] Ibid., p. 20.

[4] *The Analysis of Mind,* pp. 14-22.

[5] See however Broad's defense of the possibility of such an analysis as that of Alexander, *The Mind and Its Place in Nature,* pp. 308-311.

---

[1] *Space, Time, and Deity,* vol. II, p. 83.

[2] Ibid., vol. I, p. 25.

[3] Ibid., vol. II, pp. 81, 82.

[4] Ibid., p. 87.

The substantive aspect of Alexander's theory must now be brought out more clearly. Although mind can not "look on, as it were, from the outside and contemplate its own passing states," Alexander regards it as necessary to postulate a mind whose "connection with mental acts must be as intimate as the connection of any substance with its functions."[1] Accordingly, "every act of mind is . . . a fragment out of a larger though finite mass."[2] This mind, which cannot be directly contemplated at the human level, would become an *object* of contemplation for a super-human mind, that is, it would be seen as an object among objects, not a compresence of an enjoyed mind and a contemplated non-mental object.[3] So contemplated, mind would be seen as the highest of the finite emergents, an emergent from life. Mind can in fact be located in the body and more specifically in the "brain or some part of it."[4] "That which as experienced from the inside or enjoyed is a conscious process, is as experienced from the outside or contemplated a neural one."[5] Not all neural processes have mental aspects, but when they do "the mental process and its neural process are one and the same existence, not two existences."[6] As neural, mind may enter into causal relations with other neural processes;[7] as neural structure, mind reflects the unity or lack of unity of this structure.[8]

---

[1] Ibid., vol. I, p. 17.

[2] Ibid., p. 23.

[3] Ibid., p. 20.

[4] Ibid., p. 101.

[5] Ibid., vol. II, p. 5. C. L. Morgan's position is similar, but in regarding both the act and the content as mental, his position becomes a critical and not a new realism.

[6] Ibid., p. 9.

[7] Ibid., vol. II. p. 12.

[8] Ibid., p. 24.

There are some difficulties in reconciling the view of mind as enjoyed act, with the view of mind as neural process. While it is not impossible (though difficult) to conceive of mental acts compresent with and knowing objects that are past and future as well as present, it is difficult to see how a neural process knows a table by being alongside of it, or how a neural process in the present can be compresent with the death of Aristotle in 322 B. C. Another difficulty appears in connection with the problem of error. We have seen how in general new realism is bothered by this problem. If for Alexander mind is the awareness of objects that are independent of mind, how is error possible? Wishing to avoid the realm of subsistence, Alexander here shifts to the neural aspect of his theory. Although mind as neural complex only "selects" its content from the world, it has "interests" and can "distort" and "dislocate" this content. Most mathematical objects are "ideal selections" and not abstractions from objects,[1] but mind can also build such things as a four dimensional space which is "rather a work of art than a discovery."[2] It may even "falsify by the introduction of objects which do not belong to the thing."[3] In illusions "the mind squints at things and one thing is seen with the characters of something else."[4] "Illusions are the real world seen awry or squintingly," and such dislocation is "the mind's own work."[5] Truth, on the other hand, must be stated in terms of coherence and not in terms of correspondence: truth requires that reality "admit" the proposition,[6] but whether reality does so or not requires an appeal to the "collective mind."[1] Mind as neural process produces the illusion which mind as act simply apprehends. The relation between these two aspects of Alexander's treatment of mind is by no means clear, and it is difficult to see how in the last analysis they are identical.

---

[1] Ibid., vol. I, p. 151.

[2] Ibid., p. 163.

[3] Ibid., vol. II, p. 93.

[4] Ibid., vol. II, p. 216.

[5] Ibid., p. 216.

[6] Ibid., p. 252.

It must be admitted that a nervous process capable of such acts of apprehension, and of dislocation, and capable of the construction of four dimensional spaces is a much more remarkable neural process than any that is known to science. The conviction develops that Alexander, in endeavoring to return the contents of experience to nature while still admitting error, has invoked a dubious distinction between mental act and apprehended content, and has made a scapegoat out of the nervous system. Fortunately there are alternative methods of regarding the directly experienced world as the natural world and not merely the content of mind without invoking either diaphanous mental acts or squinting nervous systems.

The ultimate identification of the mind and the brain as found in Alexander is more prominent in the movement known as critical realism than it is in the new realism. Troubled by the problem of error in the new realism, critical realism has attempted to make a place for error by introducing between the organism and the physical world a realm of "data" or "essences," terms that at times include all content directly present and the complex phenomenon of meaning. Following closely the conventional physicist's account of the physical world, data are not regarded as constituents of this world. They may therefore be regarded as constituting a unique realm of being, as in Santayana's Platonic realm of essence, or they may be referred back to the organism for explanation, as in Sellar's account of the critical realist position. In either case a dualism has been established between the world of experience (which in many of the thinkers is regarded as mental and subjective), and the world of physical objects, which can only be regarded as a return to the conceptions and difficulties of the Galilean-Cartesian-Newtonian world-view, with a new substitute for the Cartesian mind.

---

[1] Ibid., pp. 239-241, 258.

Sellar's account is in some ways the least metaphysical and the most determined in connecting mind and the brain. For him given content is personal, psychical, and subjective.[1] It is "*in* the individual percipient."[2] Specifically, it is "an intracortical occurrence."[3] Although Sellars aims to bring the brain up to the mind and not reduce mind to the brain,[4] provided this caution is made, it may be said that "the mind is the brain as known in its functioning,"[5] and that "mental processes are brain-processes."[6] He believes that with the conception of the "brain-mind" he has enlarged "the conception of the brain to include the mind."[7] Consciousness is "a 'variant' within the brain," "a novel quality of the tensionally functioning brain."[8] "The conscious self sits in the watch-tower of the brain to guide the organism's behavior."[9]

Drake, unlike Sellars, distinguishes the mental state from the given data. Mind is not directly experienced, but the concept is necessary, so he believes, to allow the organisms to be aware of data. But in experience we perhaps get a fused and condensed content analogous to what the brain would reveal when seen from the inside. For various reasons, Drake is led to accept the position that while physics and physiology see the brain from the outside, the brain seen from within is composed of mental events. Mental events and cerebral states are therefore identical.[1] The stuff of the brain is psychic stuff; mind is "a brain regarded from the inside."[2] Mind is "a psycho-neural mechanism" whose events are "mental states;" "the bodily organ that has these states, or events, is the mind."[3]

---

[1] *Evolutionary Naturalism,* p. 32.

[2] Ibid., p. 33.

[3] Ibid., p. 39.

[4] Ibid., p. 18.

[5] Ibid., p. 300.

[6] Ibid., p. 302.

[7] Ibid., p. 315.

[8] Ibid., p. 56.

[9] Ibid., p. 318.

It may be noted in passing that Drake's metaphysics reveals the influence of both Spinoza and Leibniz. For him there is only one single universal substance, and the psychical and the physical "are simply two aspects of a single set of events."[4] The ultimate units of this substance are "psychic units."[5] Metaphysically mind is the psychic aspect of the world at the level of the brain. In addition to this single substance, Drake, like Santayana and unlike Sellars and Alexander, sets up a realm of subsistence which does not exist but which "subsists" as "merely possibilities of existence and possibilities of discourse,"[6] a realm to which "the world of existents, vast as it is, is but a speck in comparison." The realm of experience is merely the "class of essences actually 'given.'"[7]

That the movement toward the critical realist position is very widespread may be made more evident by a brief refer-

[1] *Mind and Its Place in Nature,* p. 82.

[2] Ibid., p. 93.

[3] Ibid., p. 68.

[4] Ibid., p. 90.

[5] Ibid., p. 94.

[6] Ibid., p. 198.

[7] Ibid., p. 197.

ence to two other contemporary philosophers, both of whom have been mentioned in another connection, C. D. Broad and Bertrand Russell.

Broad[1] insists that the existing object is not a constituent of the perceptual situation and cannot be inferred on logical grounds. He thus opposes the new realism both in perception and in memory. Although Broad admits that he does "not know how to define a 'mind,' "[2] his own suggestion[3] is that mind is an emergent upon the presence of two substances, one being the physical body, and one being a substance which he calls "the psychic factor," but which he admits may be physical. This psychic factor, introduced to account for psychic phenomena, may survive the death of the body, and may allow a mind to emerge when in contact with a medium's body. Whatever the merits of and need for this compound theory, and admitting that Broad introduces a new note in regarding mind as an emergent from substance and not itself a substance, there is no adequate analysis of the relation of mind when emerged to the world of nature. Although experience is defined as "a mental event owned by some mind,"[4] noises and toothaches are not regarded as states of mind or even as mind-dependent,[5] and there are places where Broad seems to make mind consist of an act rather than of apprehended content.[6] The relation of mind to experience and of experience to nature remains ambiguous, in spite of the analytical excellence of Broad's work.

Russell in his later works seems to be moving decidedly in the direction of a critical realism. Even in *The Analysis*

---

[1] *The Mind and Its Place in Nature,* p. 158.

[2] Ibid., p. 390.

[3] Ibid., pp. 535-551.

[4] Ibid., p. 376.

[5] Ibid., pp. 290, 304.

[6] Ibid., pp. 277, 305.

of Mind[1] Russell had regarded psychology as nearer to experience than physics, and in *The Analysis of Matter* all percepts have been located in the brain. Mental events are "events in a living brain, or, better, in a region combining sensitivity and the law of learned reactions to a marked extent."[2] Each percept belongs to the events "constituting the electrons in the brain."[3] Although Russell insists that physics cannot ignore quality, nor exclude percepts from the physical world, Russell gives up new realism: "we cannot . . . suppose that the external event is exactly what we see or hear; it can, at best, resemble the percept only in certain structural aspects."[4] "I should say that what the physiologist sees when he looks at a brain is part of his own brain, not part of the brain he is examining."[5] Accordingly, while Russell admits that the intrinsic quality of the "real" world may be similar to the perceived world, on his position physics can deal only with the inferred structure of the world.[6] On Russell's later position, as in critical realism, the brain has become a "brain-mind," experience has been mentalized and located in the brain, and the world which physics describes is no longer open to direct inspection.[7]

Even though no adequate discussion of such a complex movement as critical realism is possible in the present connection, a few matters of general importance may be noticed. The similarity of the general position to the Galilean-Cartesian-Newtonian world-view is evident. In place of the

---

[1] Ibid., p. 308.

[2] *Philosophy,* pp. 280, 281.

[3] *The Analysis of Matter,* p. 320.

[4] *Philosophy,* p. 294.

[5] *The Analysis of Matter,* p. 383.

[6] Ibid., p. 390.

[7] In a similar vein Morgan remarks, "all that is minded is within us. The mind never gets outside the conning-tower of vision save through projicience." (*Emergent Evolution,* pp. 50, 51).

"ivory tower" view of mind there has appeared the "ivory tower" view of the brain. What explanation can be given of the striking fact that after centuries of reaction against the world-view elaborated in the seventeenth century, the latest philosophy in time should be closest in spirit and results to this very view?

There are a number of considerations which throw some light upon this fact. The modern realisms have been nourished by the startling developments in the physical and mathematical sciences, and have not had the problems of value as their center of orientation. The world of traditional physics is a world nurtured in the metaphysics unconsciously handed down by the triumphs of Newton, and it is not surprising that those who drank of these results should enjoy the flavor of the metaphysical liquid. Lacking the handy Cartesian soul, and resenting the separation from nature which the traditional metaphysics entailed, it was natural for the new realists to objectify all experienced content and regard mind as merely an awareness of this content or a relation between separate items of this content. A mind so interpreted, however, gave no leverage for the explanation of error, left no place for constructive mental activity, and vastly complicated the world of physical reality which had been treated more simply in the traditional views. In particular, a place had to be found for mathematical systems and propositions, which although objects of awareness, were clearly not physical objects, and this led to the erection of the realm of subsistence to house the homeless universal aspects of reality. In sensing that a more adequate theory of mind was necessary to deal with the difficulties of the new realist position, critical realism, influenced by the prestige of physical science, again separated mental content and the physical object. The further tendency to locate experience, now regarded as mental, in the brain was facilitated by the enormous increase in the knowledge of neural processes, and by the growing biological cast of contemporary thought.

Two comments, however, may be made upon this result. First, the only brain that science knows is the brain as experienced, and the brain as experienced is experienced along with other factors of experience which can in no sense be reduced to states of the brain.[1] To say that a color is merely a neural process in the brain is to desert a genuine empiricism in favor of a questionable metaphysics. The brain is simply one physical object among other physical objects. If it be said that experience is the brain seen from the inside, it must be replied that this proposition is at best an assumption, and that it is assumed that there is some fundamental difference between experiencing from the "inside" and the "outside." There is, of course, another legitimate distinction that must not be confused with this illegitimate one, namely, the distinction between what is observed in one perspective from what can be observed in other perspectives. But in no perspective with which we are acquainted do we perceive colors to be brain states. To call the brain the mind is to patch up an unsatisfactory theory by recourse to a further unsatisfactory theory. Secondly, the doctrine that mind is the brain, like the process and relation theories, throws no light on the specific problems that an adequate theory of mind must meet.[2]

---

[1] Cf. Broad's discussion of radical behaviorism, to which the above criticism applies, on pp. 612-623 of *The Mind and Its Place in Nature.*

[2] Since writing the above account, I have come to realize its inadequacy both as a presentation and as a criticism of the critical realist position. There seems to me to be much that is important in the conception of essence, but nothing which cannot be dealt with in terms of the symbolic theory of meaning, and nothing which is destructive of a new realistic or objective relativistic account of the non-meaningful aspects of the experienced world.

Is there no alternative theory that will do justice to the systematic active character of mind, and to the realistic recognition that in experience we are not confronted merely with states of our mind, without being forced into a regress that leads back to the Galilean-Cartesian-Newtonian world-view, thereby losing the results of the long philosophical process from the seventeenth century to the nineteenth century? Is there no alternative theory which in avoiding philosophical blind alleys throws light upon the reflective process as immediately experienced, and has significant bearings upon the material of biology, psychology, anthropology, logic, and mathematics—a theory which while in harmony with scientific results liberates and enhances the place of mind in those cultural and moral achievements which measure the significance of human life? An affirmative answer must now be attempted.

I feel, therefore, that the more detailed consideration which critical realism deserves would necessitate no basic change in the plan or results of the present study.

III

THE FUNCTIONAL VIEW OF MIND

I. PRAGMATISM AND THE CONCEPT OF FUNCTION

IN tracing the development of the contemporary treatments of mind, it was noticed that the realistic movement vacillated between regarding mind as a relation among factors of the directly experienced world, and regarding experience as mental and as somehow imposed between the organism and the physical world which is not directly given. This second tendency, involving a regression to an earlier historical position, was primarily due to the endeavor to get enough mind to account for error without unnecessarily complicating the picture of the world. It is possible that a more adequate analysis of mind in experiential terms would keep intact the essential insight of new realism that we are not confined to our own mental states, and would avoid the regression to the "brain-mind" metaphysics of critical realism. In avoiding the metaphysical extremes of both realism and idealism, the movement of contemporary philosophy variously known as pragmatism, radical empiricism, and instrumentalism, offers an alternative approach to the problem of mind. Without failing to appreciate the systematic active features of mind, it is able to regard mind as operating in a larger non-mental environment which may be regarded as ultimately "real" and not as a mere mental modification of the world of nature. This approach involves the passage from the treatment of mind as bare relation to the treatment of mind in terms of function.

The term function, like the term substance, has a number of meanings which make a precise definition difficult. The term is often used to denote the normal mode of performance of some thing or organ, as when we speak of an engine or a heart functioning well. Closely related to this meaning, if indeed not inseparable from it, is the use of the term function to indicate the purpose which a thing fulfils or the rôle which it plays, as when we speak of the function of the heart as the distribution of blood, or the function of a particular stone on a writer's desk as that of serving as a paper-weight. There is a third meaning of function, the mathematical one, in which one term is a function of another if, when a value is assigned to the first, the second term thereby receives a specific value. In the notion of function that is to be employed in discussing the fourth type of theory of mind it is necessary to combine the elements of activity, rôle, and concomitant variation suggested in these various uses of the term. On the pragmatic position, mentality is the characteristic of experience performing in a certain functional rôle, and not a characteristic of experience in itself, experience as bare givenness being neither mental, psychical, nor subjective. The specific nature of this function will be determined as the argument progresses. Mentality, then, is similar to the status of being a paper-weight: whether a specific material object is a paper-weight or not depends upon the rôle it performs, upon the way it is used. The same object may therefore be a paper-weight at one moment and not a paper-weight at the next moment. Although functionality is undoubtedly a kind of relation, it is vastly more complicated than the bare relation of togetherness.

Historically, pragmatism shares with new realism a fundamental opposition to the metaphysics of absolute idealism. In fact, these two reactions were not in the beginning sharply separated, Peirce and James, for instance, contributing about equally to both the new realistic and pragmatic movements. The division that ultimately appeared is perhaps due to the fact that pragmatism nourished itself upon the results of biology and psychology, and showed an ever growing humanistic concern with the social activities of man, while the new realistic movement, shunning psychologism, drew its life-blood from the mathematical and physical sciences, priding itself upon an aloofness from anything that smacked of the social reformer. However antithetical the pragmatic and realistic attitudes may at first sight appear, there is some evidence of the growth of an attitude, and perhaps of a metaphysics, that is both pragmatic and realistic, the pragmatist being forced to extend his doctrine into the fields of mathematics and the philosophy of science, and the realist being forced to consider the sphere of value.

The pragmatic movement developed practically simultaneously in this country around the personalities of James and Dewey,[1] and these two thinkers have remained the complementary germanating centers of the philosophy. Speaking in general terms, it may be said that James has contributed a self-reliant individualism, a radically empirical orientation, and a metaphysical background, which in its emphasis upon the appearance of genuine novelty in the world process, opposed the static completeness of both mechanism and absolute idealism, a background in harmony with the position now known as emergent evolution. In addition to these factors, Dewey has emphasized the category of the social in working out the implications of the pragmatic attitude for ethics, education, and logic, and represents a more secular and critical attitude than is found in James. The characteristic difference in temper of James and Dewey is reflected in their analyses of mind.

---

[1] On the history of pragmatism, see Dewey's article "The Development of American Pragmatism," in vol. II of the Columbia *Studies in the History of Ideas.*

James only developed slowly in the direction of a radical empiricism. In the *Psychology* James' position is that of a dualistic realism which regards "mind knowing and thing known" as two irreducible elements.[1] "Some sort of *signal* must be given by the thing to the mind's brain, or the knowing will not occur,"[2] and in any case, "the thing remains the same whether known or not."[3] In spite of having for the criterion of mind "the pursuance of future ends, and the choice of means for their attainment,"[4] there are places in the *Psychology* where James practically identifies "mental states," "thought," "states of consciousness" with the mere fact of experience.[5] Thus the stream of experience is regarded as a stream of thought; sensations are called thoughts and are regarded as intrinsically cognitive. The bare suggestion that except for sensations "the stuff of which all our other thoughts are composed is symbolic,"[6] is not followed by any adequate treatment of the symbol. There is as yet no functional distinction of the concept from the percept, of reflective from immediate experience.

The situation has changed entirely by the time of the publication of the paper on "Does Consciousness Exist?" Here James denies what he had previously affirmed, namely, that there is any conscious stuff opposed in nature to the realm of matter, insisting that "thoughts in the concrete are made of the same stuff as things are."[1] In the further treatment of thought and mind, James really proposes two answers, not contradictory to be sure, but confused and intertwined because of James' failure to distinguish them. There is first the clear presence of a relational theory of mind,[2] the insistence that the same bit of content can now figure in a mental context, and now in a physical context, and now in both at once. The analogy which James gives is the point of intersection of two lines. The perceived room has connections with the series of events that are regarded as physical, but the same perceived room is also an event in the series of events which constitute a mind. On such a basis two or more minds may know the same thing.[3] It is in terms of this line of thought that James writes,[4] "a 'mind' or 'personal consciousness' is the name for a series of experiences run together by certain definite transitions, and an objective reality is a series of similar experiences knit by different transitions," a view which reminds one of similar statements in Hume, Mach, and Russell. This relational theory of mind is used by James to explain mind and knowledge at the perceptual level.

---

[1] *The Principles of Psychology,* vol. I, p. 218.

[2] Ibid., p. 218.

[3] Ibid., p. 219.

[4] Ibid., p. 8.

[5] Ibid., p. 186.

[6] Ibid., p. 471.

This relational explanation is at times applied by James to situations where it is inadequate, as in the case of the relation of fire to the perceived room, where he says of the room, "in the real world, fire will consume it. In your mind, you can let fire play over it without effect."[5] Here it is plain the perceived fire and the imaginatively conceived fire are not just the same content in two contexts, but rather that the conceived fire is one portion of experience acting as a symbolic substitute for other portions.

---

[1] *Essays in Radical Empiricism,* p. 37.

[2] Ibid., pp. 12, 14.

[3] Ibid., pp. 39-91, 123-136.

[4] Ibid., p. 80.

[5] Ibid., p. 14.

In the main, James himself shifts from a bare relational to a functional theory of mind in passing from the perceptual to the conceptual level of experience. On this type of explanation,[1] James no longer regards thought as equivalent to all of experience, but as concept thought is now separated from a percept by the function it performs.[2] The term intellect is practically restricted to this conceptual functioning. A concept is a percept that has become a substitute for other percepts which it leads to and which it "means." The world of concepts can rise to higher and higher levels of abstraction, and it is in the noting of the relations between these concepts that the formal studies of mathematics and logic arise.[3] The final value of concepts, however, lies in their relation to perceptual experience, in the biological rôle they play by admitting of mental experimentation in advance of overt action, and by rendering possible the analysis, recognition, and manipulation of the perceptual stream.[4] "Whenever," he says, "we intellectualize a relatively pure experience we ought to do so for the sake of redescending to the purer or more concrete level again."[5] Since the very meaning of a concept can only be stated in terms of the percepts it stands for, and the consequences for action which it suggests,[6] truth and knowledge are to be experientially described in terms of the relation of concepts to the percepts to which they lead when acted upon.[7]

---

[1] Chapters 4, 5, and 6 of *Some Problems of Philosophy* are the most important source for James' treatment of the concept.

[2] *Essays in Radical Empiricism,* pp. 61-66, 201. On p. 137 James adds that the opposition of thought and things is one of function as well as relation.

[3] *Some Problems of Philosophy,* pp. 51, 52, 67-69.

[4] Ibid., pp. 57, 63, 65.

[5] *Essays in Radical Empiricism,* p. 97.

[6] *Some Problems of Philosophy,* p. 60.

[7] *Essays in Radical Empiricism,* pp. 52-61; *The Meaning of Truth,* pp. 132, 140.

There undoubtedly remain difficulties in James' position, due in part to the attempt to treat mind and knowledge as characters of both the immediate and conceptual levels of experience. While passing towards the position of identifying mind with what we shall in the future call the symbolic portion of experience, James never attains an entirely clear-cut view of the symbol. It is not fair to make too much of an isolated statement written in the course of his own development, but any position which identifies thinking with the stream of breathing[1] cannot be said to have attained a satisfactory criterion of the mental.

Not hindered by an attempt to regard experience as such as mental or as a case of knowledge, the functional view of mind becomes much more explicit in Dewey, and to the descriptive analysis of mind in experiential terms is added an emphasis upon the social factors in the mental process. According to Dewey, experience itself is but a "natural event" with no cognitive status, but in problematic situations where the reacting organism is otherwise unable to effect an adjustment, certain portions of immediate experience function in a new rôle, that is, function symbolically,[2] or as Dewey often says, function logically. In this way portions of experience become ideas or symbols, indications of coming experiences, instruments, in short, for the reconstruction of the problematic situation. As in James' view, reflective experience arises out of and must ultimately return to the realm of immediate experience. It is the functioning of such portions of experience that gives the criterion of mind. Consciousness is the locus of such logical objects; states of consciousness are "symbols, in short."[3] Mind is "the presence and operation of meanings, ideas."[1] Mind is a "system of meanings,"[2] the "state of things in which qualitatively different feelings are not just had but are significant of objective differences."[3] As contrasted with consciousness which is "focal and transitive," mind, as "contextual and persistent," "denotes the whole system of meanings as they are embodied in the workings of organic life; consciousness . . . is the perception of actual events, whether past, contemporary or future *in* their meanings, the having of actual ideas."[4] Mind thus becomes the great tool by which man controls his world and himself, the symbolic use of the experienced world, and not a substance, not a pure act, not the bare passage of experience, and not the bare relation of experienced content to the nervous system. Mind and matter are not in any sense "static structures," but "functional characters."[5] In terms of behavior, mind may be defined as "the ability to anticipate future consequences and to respond to them as stimuli to present behavior."[6] Finally, mind so conceived may be connected with the general doctrine of emergent evolution.

---

[1] *Essays in Radical Empiricism,* p. 37.

[2] *Essays in Experimental Logic,* pp. 226-229.

[3] "The Realism of Pragmatism," *Journal of Philosophy,* 1905, vol. 2, p. 325.

There is one aspect of Dewey's treatment of mind that must receive specific mention, and that is the conception that mind is intrinsically social. Although pragmatism has always emphasized the relation of mind to action, and has always been behavioristic in the larger non-Watsonian sense of stressing the central importance of behavior for philosophy (its radical empiricism prevents it from adopting the overly simplified metaphysics of Hobbes, Watson, and Weiss), Dewey has consistently opposed the tendency to find the seat of mind in the individual brain or nervous system, and has stressed a position that might be called social behaviorism.[1] Admitting that mind is the symbolic functioning of experience, Dewey has increasingly come to regard this functioning as of social origin. While believing that "the identification of knowing and thinking with speech is wholly in the right direction,"[2] and while not doubting "that vocalization, including overt laryngeal changes, furnishes the mechanism of the greater part (possibly the whole) of thought-behavior,"[3] Dewey regards the Watsonian type of behaviorism which neglects the social situation, and which makes language a bare succession of movements in the vocal cords or movements substitutes for these movements, as giving a grossly oversimplified "subcutaneous" theory of the mental processes.[4] For Dewey the heart of language is not the expression of antecedent thought, or the bare fact of vocal cord movements, but "the establishment of coöperation in an activity in which there are partners, and in which the activity of each is modified and regulated by partnership."[5] While animals respond to stimuli of the substitutive sort, as when the hen moves away at the farmer's motion of throwing food, a child learns to react to sounds and actions "as signs of an ulterior event so that his response is to their meaning. He treats them as means to consequences. The hen's activity is ego-centric; that of the human being is participative. The latter puts himself at the standpoint of a situation in which two parties share. This is the essential peculiarity of language, or signs."[6] On this theory, animals, not having language, are not regarded as having minds. Mind is an emergent characteristic at the level of social or coöperative behavior.[1] The factor of meaning or mentality which language makes possible is then extended to other events, so that the ultimate origin of mentality is social.[2] Finally, "thinking as implicit speech . . . represents the social situation carried over into the habits of the organism. One talks to himself as a way of anticipating objective consequences (that is, consequences into which the environment enters) before they happen."[3] Dewey occasionally speaks of the "body-mind"[4] in emphasizing the relation of mentality to behavior, but such remarks should not be allowed to obscure his general doctrine that mind is the symbolic functioning of portions of experience, and that the ultimate source of all symbols, and so of mind, is the language process.

---

[1] *Experience and Nature,* p. 290.

[2] Ibid., p. 304.

[3] Ibid., p. 258.

[4] Ibid., p. 303.

[5] Ibid., p. 75.

[6] *Creative Intelligence,* pp. 39, 40.

---

[1] *Experience and Nature,* pp. 290-295.

[2] *Journal of Philosophy,* 1922, vol. 19, p. 561.

[3] Ibid., 1914, vol. 11, p. 510.

[4] Ibid., p. 509-511.

[5] *Experience and Nature,* p. 179.

[6] Ibid., pp. 177, 178.

Dewey has always thought in terms of the social categories, but it is not a very bold guess to suppose that the identification of the symbol with the language symbol represents, in part at least, the influence of G. H. Mead. According to Mead,[5] the symbol, when limited to what he calls the significant symbol, arises in those social situations where one participant can take the rôle of the other and stimulate himself as the other would stimulate him. Speech makes this possible, for in talking to others one stimulates oneself at the same time and in a way similar to the way one is stimulated when others are talking. In the use of significant symbols the user is anticipating a social situation, that is, anticipating the effects upon others of the symbols to be used. The situation may however be further "internalized," so that one becomes the other to oneself. Thinking as opposed to overt talking becomes the conversation of the self with the "generalized other." In moral conflicts, in particular, it is not unusual to be aware of an inner conversation, one member of which is the voice of the desires and urges of the individual, while the other member is the voice of the "generalized other," the social phases of the self, the social attitudes reflected in the voice of conscience. Whenever a person acts in the light of what "people" will say, the situation which Mead so acutely describes is unmistakable.

---

[1] Ibid., p. 258. Patrick has especially considered the relation of mind to the doctrine of emergent evolution in his book, *What is the Mind?*

[2] *Experience and Nature,* p. 174.

[3] Ratner, *The Philosophy of John Dewey,* p. 103.

[4] *Experience and Nature,* p. 277.

[5] "A Behavioristic Account of the Significant Symbol," *Journal of Philosophy,* 1922, vol. 19, pp. 157-163.

Markey[1] has given an interesting elaboration of the detailed mechanism involved in the Mead-Dewey theory of the symbol, but it will be convenient to postpone discussion of this, and of the philosophical bearings of the whole functional approach to mind, until after a more systematic presentation of what we may now call the symbolic theory of mind, a presentation which, while in harmony with the accounts of Mead and Dewey, in attempting to answer the question whether all symbols are language symbols, advocates a wider use of the term symbol, a use which allows of the extension of the term mind to certain animals, and to situations that are not essentially social.

II. THE SYMBOLIC THEORY OF MIND

In approaching the concept of the symbol,[2] an assumption will be made at the outset, namely, the assumption that reality appears in a plurality of perspectives with organisms as the foci or the centers of the perspectives. It is not proposed to discuss any of the philosophical problems which such an assumption involves, such problems as to whether there are other perspectives than those which are organically centered, as to whether what appears in a perspective appears in an identical form in other perspectives, as to whether the factors within a perspective are dependent upon each other for their nature and existence or whether such factors may pass unchanged in and out of a perspective, and as to whether reality consists of perspectives or whether the perspectives are perspectives of an "underlying reality." Even the problem as to whether what is given or experienced is a character of the world only in the presence of the organism will not concern us. In fact, one of the advantages of the symbolic theory of mind is that it is not dependent in any way upon the answer given to such problems. All discussions of the emergent versus the selective theory of sensa leave it untouched.

---

[1] *The Symbolic Process.*

[2] A more detailed account of the symbol, but with no express reference to the topic of mind, will be found in the *Journal of Philosophy,* 1927, vol. 24, pp. 253-262, 281-291.

That which is given at any moment is a perceptual perspective with an organism at the focus or center. The perspective called "mine" is mine only in virtue of the fact that the body called "mine" is, although only one factor among others, the focal factor of the perspective. It is the focal factor because even though at times the body is not given (as when we are said not to be self-conscious but absorbed in the "objective" world) it can easily be "recovered," and because while the body varies with the other factors of the perspective, the other factors of the perspective seem to vary in an even greater degree with changes in the body. The whole given visual world, for instance, changes with a rotation of the body, and is practically blotted out by a closing of the eyes, and is changed by eye trouble and the taking of drugs. While such changes can be validly called casual, it must be insisted that the body when given is simply one factor among the given, that it shares, as Mach would say, the same spatial field with the other given factors. Within the perspective there are discriminations to be made. In addition to my body, there also appear other organisms,[1] and other persistent factors of experience which can be approached, touched, and handled. These persistent touchable aspects of experience may be designated as physical objects. There are other aspects of experience which cannot be approached, touched, or handled, and which are less persistent and stable, such as after-images, emotions, hallucinations, kinæsthetic sensations. Instead of regarding these as mental and of a order different from the physical objects, these may be simply called non-tactual or non-physical experience.

It is possible to make plausible, though perhaps not to "prove," the assumption that certain of the objects which appear in my present perspective are themselves centers of experience or foci of perspectives. (Of course the belief in such centers of experience is not derived by reasoning. Like the child's application of the same name to widely diverse objects, there is no reason not to act as if other organisms and even other things were not centers of experience. Only as the resulting behavior of the objects fails to meet with conviction in the sphere of further experience is the sphere of this belief limited. Panpsychism and radical behaviorism represent the two extremes of the process.) If one needs to be convinced that some of the organisms which appear in "his" perspective are themselves centers of experience, the phenomena of language furnishes the most convincing examples. Suppose that I say to organism x, "take a piece of paper, start two inches from the lower left-hand corner, draw a wavy line two inches toward the diagonal corner, and around the point last drawn draw a circle two inches in diameter," and that I then take the paper and see there the same figure that I had intended to be drawn, then the conviction that the other organism is a center of experience is almost irresistible. It may be, as Broad suggests,[1] that "the position of a philosopher with no one but himself to lecture to, and no hope of an audience, would be so tragic that the human mind naturally shrinks from anticipating such a possibility," but the belief in other centers of experience besides our own plays so important a part in human life that not only the pitiable plight of the solipsistic philosopher justifies its retention.

---

[1] The expression of a body appearing in a perspective will only cause trouble if one begins with the notion that a perspective is at an instant and has no duration. But as Whitehead has maintained, such perspectives are derived by a process of abstraction from the "thick" perspectives that are given.

This perspective in which appear the focal organism, tactual objects, and non-tactual objects may be called the total self-inclusive point of view or perspective. From this perspective the "world" can only be the indefinite extension and widening of the perspective, the body must be regarded as appearing alongside of other bodies and sensible qualities, a brain is studied as directly as a color is inspected, and causal relations can only be stated in terms of experiential sequences.

Before discussing mind from this total self-inclusive point of view, it must be noticed that this perspective can be broken up into two more restricted perspectives, neither of which, it is believed, is able to furnish a satisfactory theory of mind.

The first of these may be called the partial self-inclusive point of view. To take the extreme case, imagine a person lying motionless on a couch in a dark room. Here the perspective has practically been narrowed to non-tactual given. There are some sensations of pressure, there are kinæsthetic sensations, and kaleidoscopic intangible visual experiences against a black background, but no physical objects are directly given. Imagine such a person attempting to analyze the experienced meaning of the word "dog." Upon saying or thinking the word, sensations of various sorts perhaps appear, together with various images, but little or nothing else. The meaning of the term dog in such a perspective appears simply as a non-tactual context. It is undoubtedly such situations which have given rise to the view that mind and meaning are private and subjective, and revealed by introspection.

---

[1] *The Mind and Its Place in Nature,* p. 318.

The second possible perspective may be called the self-exclusive point of view. Here observation is limited to the physical objects which appear in a perspective, and thus the content of the physical sciences appears, content that is genuinely found in experience but which is considered apart from other experienced content. If observation is further restricted to the interaction of organisms with each other and with the environment, then the perspective of the radical behaviorist is obtained. From this point of view, all that can be observed are the variations in the behavior of the organism upon the presence of stimuli, variations, to be sure, that are due as much to the state of the organism as to the nature of the stimulus. The radical behaviorist is entirely right in insisting that his position leads inevitably to a stimulus-response psychology, and that any reference to consciousness, attention, sensation, meaning, language, or mind, *in the organisms observed,* is a gratuitous assumption, and not an observation, unless these concepts are stated in terms of overt or covert bodily behavior.

There is a natural, but philosophically dangerous, tendency to go further, the tendency to interpret the total self-inclusive point of view in terms of the self-exclusive point of view. Having taken the self-exclusive point of view to other organisms, the observer of such organisms is bound to admit that another observer, observing him, the first observer, would also find nothing (except perhaps visual qualities) that could not be stated in stimulus-response terms, and the natural conclusion of this taking in imagination the behavioristic attitude to oneself, is to regard it as likely that all that is meant by thought, mind, or meaning, and even of experience itself, is some kind of bodily behavior, whether over or covert, cerebral or muscular, laryngeal or non-laryngeal. The same transference of perspective is basic to the Galilean-Cartesian-Newtonian world-view, to Watsonian behaviorism, and to critical realism, except that in the former case the content witnessed from the total self-inclusive point of view is tucked away in a substantial mind, while in the two latter cases the content is identified with some behavior of the organism or located in the brain. In brief, the whole tendency of thought exemplified in the Galilean-Cartesian-Newtonian world-view, and all of its modifications, arises in the attempt to reduce the content observable from the total self-inclusive point of view to, or to include this content in, the content observable from the self-exclusive point of view, and the only way to avoid this result is to refuse to make this reduction or inclusion, to refuse any identification of the two views.

This refusal is defensible. The most extreme objectivist must admit that he is observing something, which he calls physical objects, rats in mazes, behavior. He also sees colors connected with these physical objects. Our primary interest in the books of the behaviorist and the physicist is in the report they contain of what has been experienced. The words used are to be taken as symbols of this experience. If the writer replies, as Weiss does,[1] "for me, 'I am observing' only means 'I am responding'; and 'I am experiencing' also means 'I am responding,'" then the reader may also legitimately say that he is not so much interested in the writer's response, but would like to know something about the world. If what A observes is "in" A, and if B's observation of A is in B, and C's observation of B observing A is in C, an infinite regress is set up which destroys the whole objectivity of the perceived world, and forces the old dualism to reappear between the world as experienced and the "real" world that is not directly experienced. The attempt to include the content of one perspective in the content of another leads to the attempt to put the perspective in which the organism is found as one among other experienced factors into the organism itself. In this way a contrast between perceptual and physical space is set up.[2] The non-experienced physical world is regarded as causing changes in the non-experienced physical body, and these changes are regarded as leading to or identical with the world of experience, which is now located in the organism. In short, the Galilean-Cartesian-Newtonian world-view is present in all essentials. In refusing to admit the legitimacy of the reduction of the total self-inclusive point of view to the self-exclusive point of view, that is, by insisting upon the irreducible plurality of perspectives, one passes entirely beyond the range of assumptions of this world-view.

On the basis of the above analysis, it is advisable to return to the total self-inclusive point of view in order to frame a suitable definition of mind, particularly since this perspective does justice to both the introspectionist and behaviorist positions without being reducible to either.[1]

---

[1] "Behaviorism and Behavior," *Psychological Review*, 1924, vol. 31, p. 140.

[2] Cf. Russell, *The Analysis of Matter*, pp. 252, 253.

There is a directly experienced difference between the bare hearing of a sound or seeing of a mark and the hearing or seeing of such a word as "dog." One experience is practically or totally meaningless; the other is meaningful. In describing this difference it is undoubtedly correct to say that the meaningful word had a context different from the context of the meaningless word or sound. The former evokes bodily tensions, it tends to be followed by other words, it is felt as familiar, and in many cases, although not in all, it calls out the kind of non-tactual given that is generally called an image. But even this difference in context cannot be taken as exhausting the meaning of the term dog, which seems to "refer" or "point" beyond itself in some way. The context gives at best *the meaning*, not *what is meant*, for what is meant is a certain kind of organism which has appeared in experience, and which can be sought and pointed out if necessary. The word "dog" is a now present symbol of the now absent dogs.

Wherever there is this symbolic functioning there are things signifying and things signified, something means and something is meant. This symbolic relation is clearly acquired and involves in the last analysis the compresence of the two objects or events in experience. The phenomenon known as the substitute stimulus is so well known that a description of it is hardly necessary. It may merely be recalled that in most animals, if two stimuli objects are presented, the first of which calls out a certain response which the other does not call out, by the continual presentation of the two stimuli together, the second will finally call out the response which originally only the first called out. In such a case the second stimulus is called a substitute stimulus for the first, or the response which is thus connected with the substitute stimulus is called a conditioned response.

---

[1] It may be noted that Pavlov, while insisting that the scientist must work from the "objective" position, believes the philosopher must integrate the "objective" and the "subjective." *Lectures on Conditioned Reflexes,* p. 60.

The symbol may be regarded as an outgrowth of the substitute stimulus. It is a substitute stimulus, which while calling out in some degree the response to the original stimulus, also evokes in some form the original stimulus. In every case the symbol is a substitute stimulus, but every substitute stimulus is not a symbol.

The distinction between the substitute stimulus and the symbol may be illustrated by the case of Pavlov's celebrated dog.[1] It will be remembered that the salivary glands of the dog, originally called into activity by the presence of food in the mouth, came in time to be called into activity by the mere sight of the food, or the sight of the experimenter, or the sound of footsteps, or the sound of bells, and so forth. These have become substitute stimuli, observable from the self-exclusive point of view. There is nothing about this situation that requires the ascription of the term "mental." Suppose however that at the sound of the bell, the salivary glands not merely functioned, but that the visual image of the food was also invoked, then the bell could be regarded as a simple symbol for the food, a "reminder" of the food and not merely a substitute stimulus for it. At a more complex level, an odor may elicit familiar responses without being a symbol of the situation in which it originally appeared, and often the exact moment when the substitute stimulus becomes a symbol can be observed, that is, when the original situation is invoked or remembered. These examples may perhaps serve to illustrate the difference between the substitute stimulus and the symbol.

---

[1] Pavlov, op. cit., chap. 26.

Although any detailed explanation of this difference is in part hypothetical, it may be noticed that there is some ground upon which such an explanation can be based. If the sight of the food in the above case was itself connected with some sort of response, as is to be supposed, there is no reason why the sound of the bell, in calling out implicitly this response, should not call out the image of the food.[1] This kind of hypothesis would cause the behaviorist no difficulty since he could identify the image and the implicit movement, nor would it clash with any position which regarded the experienced object as being in part due to the organism, for in such a case the movement would simply be part of the conditions necessary for the emergence of certain sensa. With a little ingenuity such a hypothesis could even be reconciled with the realistic doctrine that sensa are selected by the organism, not produced by it, for in this case the movement of the organism simply serves as one of the conditions which determine this selection. Hunter would not wish to use the word "image" (and even on the present account all symbols do not involve imagery), but his position that a symbolic response is a "substitute process which can be reinstated by the organism, but *only* where associative traces of the original process persist in the integration,"[2] suggests a possible physiological basis for the symbolic process, and if the traces of the original process are present in sufficient degree, there seems to be no objection to describing this as a rearousal in some form of the otherwise now absent stimulus.

Any portion of experience, then, that has become a substitute for and a reminder of some other portion of experience may be called a symbol. Or more precisely, "a symbol is any given or experienced substitute stimulus that leads to a reinstatement of the original stimulus in a form that is only observable from the self-inclusive point of view."

---

[1] See M. F. Washburn, *Movement and Mental Image,* chapters 3 and 4, and pp. 30, 31 in particular.

[2] "The Symbolic Process," *Psychological Review,* 1924, vol. 31, p. 488.

The school of thinkers led by Mead and Dewey would probably regard this characterization of the symbol as defective for two reasons: as not sufficiently stressing behavior, and as employing the term symbol below the level of language symbols. As for the first point, the present theory recognizes the closest relation of the symbol to the organic process of adjustment, but it does not believe it is advisable to use the term response to cover all experienced qualities,[1] nor that it is necessary to attempt to identify such qualities with segments of physiological behavior.[2] The only behaviorism this present theory would not harmonize with is the oversimplified Watsonian behaviorism. As for the second point, it is doubtful if it is necessary to reserve the term symbol for socially acquired symbols. In this connection we may glance at Markey's treatment of the symbol.

For Markey,[3] symbols arise only in social-vocal situations, the material for the symbolic process being found in the vocal sounds of the child that have been circularly conditioned, that is, sounds the hearing of which has become a substitute stimulus for the uttering of the sound. Given such conditioned circular responses, the voice of the mother, when it repeats a sound of the child that has been circularly conditioned, tends to call out the utterance of the sound by the child. Further, by a process of substitution, the presence of the mother and even objects near her, may become substitutes for the mother's voice, and thereby come to call out the vocal response of the child. This vocal response, finally, comes to stimulate the child as the previous presence of the mother and her voice stimulated the child, and such a vocal response, performing the stimulating rôle of another person, but differentiated from the other person's voice by the child's own original responses to his own voice, is a symbol. Symbols arise, accordingly, out of the social-vocal situation. Thinking, which at first sight does not seem to be necessarily social, is explained as the internalization of the social act of conversation.[1]

---

[1] As Markey does, *The Symbolic Process*, p. 33 n.

[2] Neither Mead nor Dewey is guilty of such a reduction, I believe, but to avoid confusion the sense in which they use the term must be carefully noted. For Dewey, behavior is "more than physical." (*Journal of Philosophy*, 1914, vol. 11, p. 510) On Dewey's use of behavior, see Ratner, *The Philosophy of John Dewey*, p. 100.

[3] Op. cit., particularly chapters 3 and 9.

Now there can be no doubt that this interesting account gives the way in which the majority of the important and fairly permanent symbols have arisen, and that the Mead-Dewey-Markey account of the symbol is one of the most significant analyses of certain aspects of mind in contemporary thought. But a number of considerations make it advisable to hesitate in identifying all symbols with language symbols growing out of the social-vocal situation. There is some doubt as to whether either the vocal or the social aspect is indispensable. First as to the vocal aspect. One implication of this theory would be that a family of deaf-mutes, not especially trained by normal individuals, would develop no symbols and no mental phenomena. This seems highly questionable, but since this is open to some kind of determination by the study of deaf-mutes, and some check afforded by experimental evidence, no dogmatic statement one way or another is as yet desirable. Secondly, while the older theory of a gestural language preceding vocal language is passing into the discard, it nevertheless does not seem likely that all gestural symbols are derived from vocal symbols. Thirdly, although writing is undoubtedly very young in comparison to spoken language, the fact that writing was pictographic in origin and only later became phonetic suggests that there may have been subsidiary forms of symbolism, visual in this case, that have been supplementary to, and not derived from, the vocal symbolism.

---

[1] This particular point, found also in Mead and Dewey, is developed in Grace de Laguna's *Speech: Its Function and Development,* pp. 341-343.

As for the social aspects of the symbolic process, the question is as to whether another person is absolutely necessary for the symbol, even the vocal symbol, to develop. Admitting that the case is an extreme one, echoic phenomena might occasionally take the place of the other person's voice. Then there are certainly sounds in nature which could perform this function, and the onomatopoetic aspects of language would be explained on this assumption. There is nothing in the hypothetical physiological basis of the symbolic process which make it impossible to conceive of the conditioned circular vocal responses being directly conditioned by non-auditory aspects of the world.[1] Could not a stimulus S by occurring along with the (perhaps physiological) stimulus which calls out a vocal response become a substitute for the original stimulus and thereby call out the vocal response directly? Once this is done S^1 could become a substitute for S, and even a symbol of it. Finally, without any vocal response at all, it seems possible that of two objects experienced together, one could become not merely a substitute for the other, but could lead to the rearousal of sensory processes connected with the other object and in this way become a symbol of it.

---

[1] Allport, whose work has been basic to all later theories of language, does not hesitate to make use of such an assumption in explaining how language first arose. See his *Social Psychology,* pp. 194, 195.

If to these considerations be added the further considerations that the individual and society are correlative, that the possibility of the symbol must be found in the mechanisms of the individual, and that no sharp line is to be set up in the evolution of intelligence from the animal world to man, it may be seen why it is felt to be desirable to adopt a view of the symbol that would not restrict the symbol to socially acquired symbols, but which would be able to combine Hunter's use of the term with that of Markey's. Fortunately no difference is made in either case in the general theory of mind, the question being simply one as to the range of the application of the term "symbol." However used, the important thing about the symbolic process is that it allows the appearance in some perspective of objects and situations not present in the direct sense in which they were originally present.[1] We have preferred to regard the symbolic process as having many roots, and as extending into the animal world, the social-vocal situation giving the most effective and important development of a process that need not be either vocal or social, but which attains its highest development in the social-vocal situation. Should it seem advisable at a later time to limit the extension of the term symbol to socially acquired symbols, the range of mind on the symbolic theory would be correspondingly limited but the theory as to the nature of mind would not be changed.

On the symbolic theory the criterion of mind is the presence of symbols in a perspective. Mentality is the symbolic functioning of events. Speaking in the most general form, mind is the locus of symbols. If it is preferred to adopt Dewey's distinction between mind and consciousness, it may be said that consciousness denotes the present functioning of symbols, while the phrase "a mind" denotes the entire systematic symbolic repertoire of an organism.

---

[1] Thus Markey writes, op. cit., p. 138, "within a symbolizing behavior system there is one part of the integrated behavior *which stimulates and presents absent situations, past events, and possible future events*—the whole range of the universe for which adequate symbols are at hand." (Italics mine.)

III. SOME IMPLICATIONS OF THE SYMBOLIC THEORY OF MIND

There now remains the need of briefly considering the symbolic theory of mind in terms of the philosophical, psychological, anthropological, and logico-mathematical criteria which were suggested as tests for any adequate theory of mind. In this consideration, certain implications of the theory will be suggested.

To begin with the philosophical implications, it is evident that the symbolic theory of mind combines to some degree the central claims of the rival theories of mind: with the idealists it admits that the content of mind is the world, and that mind is an active systematic process; with the new realists that the world as experienced is not a collection of mental states or in a mind; with the critical realists (and here the divergence is greatest) that the active, unitary, "substantial" aspect of mind is dependent upon the unity and activity of the organism.

The theory also avoids any dualism between mind and the experienced world, or between mind and the body. Since any experienced content may function symbolically, any portion of the experienced world may put on or take off the status of being mental by functioning or ceasing to function symbolically. To a large degree the symbolic theory of mind is metaphysically neutral and it could be harmonized with any one of the typical philosophical theories. On this view, a dualism between mind and the world could only arise if the world of experience (and so of mind) is regarded as different in kind from the world of "reality."

The symbolic theory of mind is able to deal with truth or knowledge,[1] and the converse, error, provided that truth and error are regarded as describable in experiential terms, and not as requiring a relation of experience to a non-experienced reality. By the compounding of symbols, which allow of the compounding of meanings, predictions as to further experiences may be made, and when the experience as predicted is the experience obtained, the prediction is said to be true and a case of knowledge. When the completing experience is not obtained as predicted, then the prediction is said to be false. Just as experience is not as such mental, so it is not as such true or false or a case of knowledge, all of which depend upon the symbolic functioning of events.

As for the relation of the present theory to the positions of the psychologist, the present position, which in psychological terms may be called a neo-functionalism, utilizes the material of both the existential and the behavioristic psychologist, while rejecting the claim of either to give by itself an adequate account of mind. Mind is regarded here neither as the content of experience considered in its dependence upon the nervous system, nor as reducible to physiological behavior. Nevertheless, it may be agreed with the existential psychologist that a complete description of meaning requires the inclusion of material not observable from the behaviorist's perspective, and with the behaviorist that meaning is not statable without reference to the behavior that is involved in the substitution of stimuli.

Because of the fact that every symbol involves behavior, the employment of symbols can effect even the glandular processes of the organism, and so the present theory is capable of being extended to the phenomena of auto-suggestion and hypnosis. Time does not permit a consideration of the bearings of the theory on such topics as the dissociation of personality, and mental disease.

---

[1] The topics of truth and knowledge are treated from the symbolic point of view in two articles in the *Monist,* 1928, vol. 38: "The Prediction Theory of Truth," pp. 386-401; "Neo-Pragmatism and the Ways of Knowing," pp. 494-510.

As for biological considerations, the symbolic theory of mind allows both for the tracing of the development of intelligence in the animal world and the exhibition of the way in which mind has maintained itself in the selective process of evolution. As for the first point, with the present definition of the symbol, it is possible to regard the rudiments of mind as appearing in animals below the human level. While no agreement in the interpretation of the delayed reaction experiments can be claimed,[1] there is much in such experiments that justifies Hunter's use of the concept of the symbol in the interpretation of such experiments. At the same time it may be admitted that many animals show no behavior that seems to require the introduction of the symbol, and that wherever mind exists in the animal world, it is probably below the level of the language-symbol. So that while bridging the gap between animal behavior and the human mind, the theory is able to do full justice to the unique development and capacity of the human mind.

Indeed, man's place in the world of living beings is due to the extensive development of the symbolic process which human society has permitted. While even the simple symbol may be of some biological value, since the rearousal of the original stimulus presents new stimulating features in addition to those of the substitute stimulus,[2] the extensive development of language symbols makes possible the funding of many experiences, past and present, and thus makes possible the effective organization of social resources in dealing with social and individual problems. Symbols may be regarded as mental antennae, the farthest present extension of the process which has passed from the dealing with objects by touch, to the dealing with distant objects by the development of the distance receptors, and finally to the dealing with absent objects through the functioning of their present symbolic substitutes. In this sense, mind is the furthest extension of the learning process, the most effective form of making adjustments which the evolutionary process has produced and maintained.[1] The relation of the symbolic process with the evolutionary process is intimate and illuminating.

---

[1] Markey is willing to admit that the symbolic process is built upon the delayed reaction behavior, but does not wish to use the concept of the symbol to explain such behavior. See op. cit., chapter 8.

[2] Perhaps what Pavlov says of substitute stimuli may be true here also: "the sum of the stimuli acts more strongly than they do separately" (op. cit., p. 54).

Anthropologically, the mind of primitive man is seen to differ from the mind of modern man in the number, range, abstractness, permanence, and definiteness of the symbols, and in their coherence and interconnectedness. To ascribe the characters of "prelogical" and "mystical" mentality to early man is to be in danger of exaggerating the differences of primitive and modern mentality. It is true that primitive man has no adequate technique for separating the causal and non-causal sequences of experience, that he often confuses and identifies the symbol and that which is symbolized, and that objects often have different meanings than they have for us, but to some degree all of these factors continue to appear in most contemporary minds. The recognition of the differences between primitive and modern mentality should not obscure the fact that these differences are earlier and later stages of one cultural process. Although there are perhaps native differences between "races" and individuals in the capacity for symbolic attainment—differences perhaps in a process of evolutionary selection—the main differences between modern and primitive man are undoubtedly expressible in terms of social heredity rather than in terms of "mental" ability. Also, the recognition of the individual phases of the symbolic process (the recognition that however important the social factors in this development, the symbol requires contact by the individual with actual objects)[1] prevents the excessive emphasis upon the "collective representation" aspect of primitive symbols. Here, as elsewhere, the individual and society are correlative.

---

[1] Cf. Thurstone, *The Nature of Intelligence,* chap. 5. The social function of language symbols is well developed in the previously mentioned work of Grace de Laguna.

Only a word can be said about the logical and mathematical implications of the symbolic theory,[2] and even then only the formal aspects of these studies will be referred to. It has already been noticed that the tendency of the realistic philosophy was to withdraw mathematical and logical propositions from the realm of mind and the realm of matter, and to regard them as inhabiting a unique realm of being, the realm of subsistence. The symbolic theory of mind is able to avoid such a realm, and to take account of the genuine creativity of the mathematician and logician, by regarding logic and mathematics as a development of, and a specialization within, the symbolic process. By this orientation, the formal studies are linked with and share all the benefits of the reflective process. Neglecting the question as to the origin of formal propositions and the question of their truth or application outside of the formal system, it may be said that the distinguishing characteristic of the mathematical and logical types of reflection lies in the generality of the symbols employed (the symbols standing for any other objects that meet the specified requirements), in the use of symbols defined entirely in terms of other symbols (such as the square root of minus one, and the point at infinity), and in the restriction of interest to the interconnection or consistency of the symbols (which is expressed in the postulational deductive form which the formal studies ultimately assume). Symbols have relations to each other as well as to non-symbolic events, and the formal disciplines study in its most general form this interrelatedness and mutual dependence of symbols. By being connected with the symbolic process, no diremption of mathematical or logical objects from either mind or nature is necessary. This suggestion is fragmentary, to be sure, but it indicates that the symbolic theory offers the possibility of accomplishing that which has constituted a signal weakness of most empirical philosophies, namely, an empirical approach to the domains of formal logic and pure mathematics.

---

[1] Goldenweiser criticizes the neglect of the technological activities of the individual in the theories of Levy-Bruhl and Durkheim on pp. 380-389 of his *Early Civilization;* and criticizes the doctrine of collective representations in the *American Anthropologist,* 1911, vol. 13, pp. 121-130.

[2] This topic is treated in some detail in an article, "The Relation of Formal to Instrumental Logic," contributed to the volume, *Essays in Philosophy,* The Open Court Publishing Co., 1929.

Finally, the symbolic theory of mind is not a "silly" theory. It can do full justice to mind—it "saves the appearances" of mind—without unduly exaggerating the place or nature of mind. It makes intelligible popular expressions which refer to one mind as better than another, which speak of the development of a mind, of the dissociation of a mind, of the unity of the mind, of the activity of the mind, and the like. Without reducing mind to physiological behavior, the theory explains the close relation of mind to behavior. By the capacity of mind to influence behavior, in the fact that the rational being is able by the operation of symbols to determine in part his own behavior, is to be found the genuine source and meaning of human freedom, of moral value, and of moral responsibility; in the anticipation and partial control of future consequences which the operation of symbols allows, is to be found that earnest stimulation to coöperative endeavor which makes man the individual and man collectively a determiner of destiny and not a mere puppet in the hands of Chance. In the socially oriented life of mind is found the factor which gives unique value to human striving, and which furnishes the only sound basis of legitimate hope.

Charles W. Morris.

MATHEMATICAL LECTURES[1]

I

CONSISTENCY IN MATHEMATICS

THERE are two circumstances which formally evoke the danger of contradiction within the system of mathematical propositions, because they prevent these propositions from being significant statements, in the sense that we know what we mean by asking whether they are true or false. The one circumstance, as Brouwer first made clear, is the unlimited application of the terms "all" and "any" to a field of mere possibilities which is open to infinity; the other is the leveling process which mathematics blindly performed on the Russell types. Especially with regard to the second point, mathematics manifests its full participation in the servile revolt of the positive sciences against philosophy, the revolt of the anti-spiritual mind, with its democratic leveling process, against the spiritual mind and its hierarchic structure, which changed the question: "What is your intrinsic nature and what does this nature bring forth?" into the other: "What can you be used for? What profit do you yield when you are made to play your part in the process of production standardized by such and such axioms?" Brouwer's intuitionistic mathematics represents the restoration of mind to its old and sacred rights. Hilbert's formalized mathematics, however, undertakes to show that the opposite party, which indeed sinks far below the mind when it demands that its overflowing wealth of "results" be accepted as literally true, is ultimately right in spite of all—ultimately, however, meaning: before a transcendental forum which we realize symbolically. In mathematics the inquiry into the genuineness or non-genuineness of the inner working of our entire western culture urges towards a more rigorous decision than can be attained in the other hazier fields of knowledge.

---

[1] Lectures delivered at the Rice Institute, May 20, 22, and 23, 1929, by Hermann Weyl, Professor of Mathematics at the Technische Hochschule in Zürich.

With regard to the first point, the usage of the terms "all" and "any," I think one does not hit quite the right spot by referring to the validity or invalidity of the principle of the excluded middle. As you know, the point in question is the confrontation of the two assertions: "Being given a set M of objects, there exists an element in M with the property $\mathfrak{A}$" and "All elements of the set M have the property non-$\mathfrak{A}$." But the stress is not on the fact that two assertions are confronted one of which occurs as the negation of the other, but on the fact that these assertions involve the terms "there exists" and "all." Moreover, it is incorrect to describe the intuitionistic point of view by saying that the *tertium non datur* applies or does not apply in the case referred to, according as M is a finite or an infinite set. The issue does not lie in the distinction between finite and infinite, but it depends on whether M is given as an aggregate of objects which are individually exhibited, one by one (and is therefore indeed finite), or not. If several pieces of chalk lie in front of me, the assertion: "All these pieces are white" is merely an abbreviation for the assertion: "This piece is white, and this piece is white, and . . . " (while I exhibit them one by one); likewise "There exists among them one red one" is an abbreviated expression for: "This one is red, or this one is red, or · · · ." But such an interpretation is possible only for sets the elements of which are exhibited. If, in opposition to the given example, we take the sequence of natural numbers 1, 2, 3, · · · and consider an assertion such as "All numbers are even," the analogous interpretation leads to an infinite logical product (I put the logical "and" and "or" into analogy with the arithmetical × and +): 1 is even, and 2 is even, and 3 is even, · · · . But this obviously has no meaning. Wherever a general proposition of this kind occurs, it has a hypothetical meaning, it assures that if you are given any definite number, for example 18, you are certain of the correctness of the judgment that 18 is an even number. It is evidently impossible and without meaning to negate such a hypothetical proposition. This fact that the negation cannot be carried out, and not the invalidity of the *tertium non datur*, is the point on which the matter hinges. The formal negation of our general judgment: "there is an odd number" would be equivalent to an infinite logical sum: 1 is odd, or 2 is odd, or · · · ; it gains significance only with a view to the explicit construction of an individual definite number, for example 17, which is established to be odd. I have therefore called the existential proposition an abstract of judgment. If knowledge is a valuable treasure, I compare this abstract to a paper which informs us of the existence of a treasure but without disclosing where it lies. With regard to the transition from finite logical sums and products to infinite ones, matters stand much the same as in the domain of arithmetical operations, where the definition of a finite sum does not *a priori* determine the meaning of an infinite sum.

I do not want to go into too much discussion to convert you to this opinion of Brouwer's. It is entirely a matter of reflection (*Besinnung*),[1] which has nothing to do with any epistemological, or perhaps even metaphysical theories, nor indeed with any arbitrarily declared mathematical axioms and their technical manipulation. Everybody will admit its truth provided he understands it.

In the development of arithmetic, we can perhaps distinguish four stages with regard to the part played by the infinite. To the first stage belongs a concrete individual judgment like $2+3=5$. To the second, for example, the following judgment of hypothetical generality: "If m and n are any two concretely given number signs, then $m+n=n+m$." I can comprehend the meaning of this proposition and convince myself intuitively of its correctness, without "generating" any other number beside the two concretely given ones m and n; it is not even necessary to form the number $m+n$. In the third stage, the actually occurring number signs are imbedded into the sequence of all *possible* numbers, which originates by means of a generating process according to the principle that from a given number a new one, the following one, can always be formed by addition of the number 1. Here the existent is projected into the background of the *possible*, the background of a manifold of possibilities which is produced and ordered according to a fixed process but is open into infinity. To this point of view corresponds the method of definition and conclusion by means of complete induction. I cannot conceive of a grosser misunderstanding than that of making the legitimacy of this procedure which refers to the *possible* depend, as Russell does, on the actual existence of infinitely many objects in the real world. I believe that here we strike the root of the mathematical method in general: the *a priori* construction of the *possible* in opposition to the *a posteriori* description of what is actually given. In the fourth place, however, and this is where according to Brouwer the fault of mathematics begins, the theory of sets declared the sequence of natural numbers, a sequence open into infinity, to be a closed complete aggregate of elements existing in themselves. Take the definition: "n is even or odd according as there does or does not exist a number x such that $n=2x$." Whoever accepts this definition which appeals to the infinite totality of numbers as having a meaning, has passed into another sphere; for him the number system has become a realm of absolute existences which is "not of this world" and of which our consciousness catches a glimpse only here and there.

---

[1] L. E. J. Brouwer, "*Intuitionistische Betrachtungen über den Formalismus,*" Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1928, pp. 48-52. A bibliography is included.

The second weak spot in the body of mathematics is due to the objectivation of properties, as Russell emphatically pointed out. As long as we deal with the natural numbers as objects of investigation, single definite properties such as being even, being a prime, etc., occur. The new step is made, when a judgment like "18 is even" is no longer subsumed under the propositional form, "x is even" by the substitution $x=18$, but under the form "x has the property y" with two empty places, x and y. Properties have hereby become objects of a different order, to which the primary objects, the numbers, can entertain the relation ϵ of "participation," μέθεξις, as Plato says. This step becomes dangerous only when the terms "all" and "any" are applied without restriction to this new realm whose objects are "the possible properties." In this sense the second step is already included in the first. It is known that the constructive generation of properties establishes a hierarchy of types, the Russell types, the neglect of which leads to striking contradictions. An objectivated property is usually called a *set* in mathematics; and the real number of analysis is essentially equivalent to the notion "set of natural numbers." Therefore this second step is of decisive importance for the foundation of the theory of the continuum as the field of all possible real numbers. Usually the word "set" when used in opposition to "property" includes this further convention: the sets corresponding to the properties $\mathfrak{A}$ and $\mathfrak{A}'$ shall be considered equal if every object with the property $\mathfrak{A}$ also has the property $\mathfrak{A}'$, and vice versa. I shall here, however, leave aside the difficulties connected with carrying this identity principle into effect in a rigorous manner.

The intention of the Hilbert proof theory is to atone by an act performed once for all for the continual titanic offences which mathematics and all mathematicians have committed and will still commit against mind, against the principle of evidence; and this act consists of gaining the insight that mathematics, if it is not true, is at least consistent. Mathematics, as we saw, abounds in propositions that are not really significant judgments. But we must abstract from the content of *all* its propositions and consider only their formal structure when we intend to show that they involve no contradictions. Thus mathematics becomes, in Hilbert's theory, a game with signs and formulas; the formulas, which consist of signs, have no meaning which they wish to convey, but they are the material of the game of demonstration: according to the rules of the game new formulas are constructed from those already at hand. The formulas that one starts with are the axioms. Among the signs the negation $\backsim$ occurs. We would have a contradiction if of two proof games which both start from the axioms and are played according to the rules of the game, one ended up with the formula $\mathfrak{b}$ and the other with the contrary one $\backsim\mathfrak{b}$. The point is to gain the insight that this can never occur, and this is an act of cognition and not play. Exactly as we can convince ourselves that in a correctly played game of chess a position with nine white pawns can never occur. The insight into the consistency of the game of mathematics has to be attained in the same direct way as that referring to the chess game. Here the consideration proceeds as follows: in the initial position there are eight white pawns. According to the rules a move may decrease, but cannot increase, the number of pawns. Ergo · · · This ergo stands for the conclusion by complete induction which follows the concrete given chess game, move by move. It is self-evident for Hilbert that the considerations by means of which the consistency proof is given in "metamathematics" are throughout endowed with the finite character postulated by Brouwer. This intuitive thinking in terms of content matter is based on evidence and not on axioms; it is conveyed by means of language which is necessarily always an uncertain tool of communication. On the other hand, mathematics itself has no need of any language, since its formulas mean nothing and convey nothing.

But why go beyond the bounds of significant judgments, since what lies beyond is totally empty and cognition can gain nothing from it? A possible answer to this question appears to be that which assigns to the ideal judgments a part similar to that played by paper money in economy: it does not add new values to the real ones but it makes their handling easier. Whenever, in Hilbert's formalized mathematics a proof yields a final formula which admits interpretation as a significant judgment, this judgment is true. But I hardly think that this purely technical employment of the formulas for the deduction of significant propositions would sufficiently justify the method. Still this is not a controversy of principle but only a question of economy. Hilbert himself gives the somewhat obscure answer that the infinite plays the part of an idea in the Kantian sense, namely that it supplements the concrete in the sense of totality. I hope I am in agreement with Hilbert when I interpret this as analogous to the construction by which I imbed the objects which are actually given to me in my consciousness into the totality of an objective world which comprises many things that are not immediately present to my mind. From the point of view of pure consciousness, it is also here not at all easy to understand what this supplementation really means. Certain epistemological schools would like to interpret it as being only a technical artifice which enables one to find one's way about more easily in one's own consciousness. But there are enough people, and I belong to them, who have the firm belief that somehow the reality of the "you" and of the exterior world embodies a higher truth than this solipsistic point of view. Theoretical physics justifies and completes the construction of that intersubjective world we build up in our natural life. The conditions which prevail here by no means correspond to Brouwer's ideal of a science. An individual assertion, an individual physical law has no meaning that can be realized by intuition and verified by experience. Only the theoretical system as a whole is capable of confrontation with experience. And it holds good if *concordance* prevails, that is, if on the basis of our theories, all indirect determinations of the same physical quantity lead to the same result.

An example will make clear what I mean. Let us consider a definite oscillation of a pendulum; and let us assume that its period can be observed directly with an error less than 0.1 second, so that periods of oscillation which are described by the theoretical physicist as differing by less than 0.1 second are actually equal, i. e., equal for our direct perception. Still there is a simple way of increasing the exactness of observation a hundredfold: one waits until 100 oscillations have taken place and then divides the observed period by 100. But this indirect determination is dependent on a certain hypothesis, namely on the hypothesis that all oscillations last the same length of time. This can of course be tested by the direct observation with an exactness of 0.1 second. However, if the theory is used in the indicated manner, not this is meant, but instead that the periods are absolutely equal, or equal with a hundredfold precision. This assumption, just as well as the assertion with regard to the period of an individual oscillation to which it leads, is without meaning for the intuitionist who respects the limits of intuitive exactness. Still it is possible to test the hypothesis in a certain sense: one finds that the period of duration of m successive oscillations is to that of n oscillations as m is to n, where m and n are large numbers (for the test we arbitrarily choose several series of oscillations). If I interpret Hilbert correctly, an analogous situation is already prevailing in pure mathematics.

The formalist who abides by his principles must leave the question unanswered why he chooses just *these* axioms for the starting-point of his proof game. Also his interest in the fact that no contradiction occurs can hardly be justified or can at most be justified by the following remark. If two games lead to the formulas $\mathfrak{b}$ and $\backsim\mathfrak{b}$, then, if $\mathfrak{a}$ is an arbitrary given formula, it is possible to obtain the formula $\mathfrak{a}$ by two additional moves, as final result. It is consequently *a priori* certain that one can prove any arbitrary formula $\mathfrak{a}$ and one has a simple fixed rule according to which to do it. In this case the game would be tedious; still it would only be tedious if I knew the contradiction. If, however, we consider this game of formulas as a symbolic expression of a theory about the world, consistency is involved in the above described concordance. Thus we get a more satisfactory answer: only a consistent theory can lead to concordant results when it is applied to experience; the consistency is that part of the concordance which refers only to the theory itself, the part in which the sphere of what is sensually given is not yet touched. It is the task of the mathematician to see that the theories of the concrete sciences satisfy this condition *sine qua non* of being formally definite and consistent.

The development of science has shown clearly that different theoretical constructions of the world satisfy the postulate of concordance. The decision between the theories which compete in this manner is practically cogent for every open-minded scientist, yet it is hard to say precisely what brings it about. On the other hand it is scarcely up to us, the mathematicians and physicists, to account for this question; for in this respect we are at the mercy of the decisions cast in the history of mind, destitute of that ultimate insight that Brouwer postulates. What truth means in physical theories is a philosophical or epistemological problem rather than a physical one.

After these general remarks, I should like to go into a more detailed discussion of the structure of Hilbert's formalized mathematics.[1] As long as the transfinite is excluded, only two kinds of signs occur in the formulas; the *constants* like 1, 2; and the *operations*. Logical operations are $\backsim$ (negation), & (and), $\vee$ (or), $\rightarrow$ (implies); the first one is one-membered, the others are two-membered. The two-membered operations $=$ (is equal to) and ϵ (is element of) may be considered as logico-arithmetical ones. The operations σ (generates out of a the natural number following a)

---

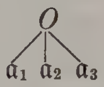
[1] This description follows, however, more closely a paper of J. v. Neumann, (*Math. Zeitschrift*, vol. 26, 1927, page 1) than Hilbert's own formal system.

and Z ($Z\mathfrak{a}$ read: $\mathfrak{a}$ is a natural number) are purely arithmetical one-membered operations. But I should explain the sense in which I speak throughout of operations and not of relations and properties besides. →, for example, stands for the operation which generates the judgment: "$\mathfrak{a}$ implies $\mathfrak{b}$" from the two judgments $\mathfrak{a}$ and $\mathfrak{b}$; Z is the operation which generates the assertion $Z\mathfrak{a}$: "$\mathfrak{a}$ is a natural number" from $\mathfrak{a}$. These remarks are of course merely explanatory and are intended to recall the correspondence between the formulas of our formalized mathematics and certain propositions of ordinary mathematics which are meant as actual assertions of something.

If we include the infinite, two new kinds of signs become necessary: *variables* like x, y, and *integrations*. By means of the integration Σ_x the assertion $\Sigma_x\mathfrak{A}(x)$: "There is an x for which $\mathfrak{A}(x)$ holds," is obtained from the proposition $\mathfrak{A}(x)$ with one variable x. It is distinguished from the operations by the fact that it contains an arbitrary variable x as an index and "ties up" this variable in the formula standing behind it, i.e., deprives it of its capability of being substituted, exactly as though a definite thing, a constant, had been substituted for this variable. We are now in a position to describe in general what a *formula* is. Let it be written in the form of a genealogical tree such that an operation appears as the father of the terms on which it works. Thus a 1, 2, · · · membered operation is always followed by 1, 2, · · · signs respectively; the immediate progeny of an integration consists of a single sign, while a constant or a variable is always a last member without descendants. At the head of the genealogical tree we find a sign of integration or operation, and all its branches end with constants or variables. We can describe the same thing inductively as follows: An individual constant or variable in itself constitutes a basic formula. Out of these basic formulas, we construct derived formulas according to the two following principles: (1) If, for example, three formulas $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$, are given, and O is a definite three-membered operation, a new formula is obtained by appending $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$, separately under the sign O. (2) If a formula $\mathfrak{a}$ is given, and if J_x is an integration which carries a certain variable x as an index, a new formula is obtained by hanging $\mathfrak{a}$ under the sign J_x. To make a homogeneous description possible, we thus, for example, write $\overset{=}{\mathfrak{a}\diagup\quad\diagdown\mathfrak{b}}$ instead of $\mathfrak{a}=\mathfrak{b}$. But afterwards we may return to the habitual way of writing so as to prevent the formulas from having too strange an aspect. Let us do this, and also replace the operational symbol σ by the symbol $+1$ which shall be put after the formula to which it refers ($\mathfrak{a}+1$ instead of $\overset{\sigma}{\mathfrak{a}}$). Furthermore I have to describe the process of substitution. $\mathfrak{A}$ or $\mathfrak{A}(x)$ may be, as we always assume in the following, an arbitrary formula in which at most *one* variable x occurs free (not tied up), and $\mathfrak{b}$ (or $\mathfrak{c}$) a formula without a free variable, a so-called normal formula, then x shall be replaced by the entire formula $\mathfrak{b}$ everywhere in $\mathfrak{a}$ where it occurs *free*. The process of substitution which is thus intuitively described, again produces a formula; this is the formula we have in mind when we use the abbreviation $\mathfrak{A}(\mathfrak{b})$. The substitution rule would turn out to be more complicated if $\mathfrak{A}$ contained several variables free and if free variables also occurred in $\mathfrak{b}$. For example, let $\int_x$ stand for integration with respect to x from 0 to 1; then it is permissible to replace y in the correct formula

$$\int_x xy = \tfrac{1}{2}y$$

by a constant or a function which does not contain the



variable x. On the other hand, nonsense results when y is replaced by a quantity containing x, for example by x itself: $\int_x xx = \frac{1}{2}x$ instead of the correct result $\int_x xx = \frac{1}{3}$. However, we avoid these complications by the restriction mentioned above.

German letters are used in general descriptions for means of communication; they belong to the language and are not signs in the same sense as 1 or x: men at our game of mathematics. In the course of the development of mathematics, new signs can continually be introduced; but of course we must always make the accompanying remark that the sign is a constant, a variable, a 1, or 2, or 3, · · · membered operation, or an integration.

Now for the axioms! First the finite logical axioms such as

(⁑)
$$\mathfrak{b}\rightarrow(\mathfrak{c}\rightarrow\mathfrak{b}).$$

It states: if you have any two definite formulas $\mathfrak{b}$ and $\mathfrak{c}$ without free variables, put them together to the formula $\mathfrak{b}\rightarrow(\mathfrak{c}\rightarrow\mathfrak{b})$; you may then use this formula as an axiom. Thus (⁑) is not an axiom itself, but a general rule for formation of axioms, an inexhaustible source of axioms. It is not necessary here to enumerate the few finite logical axiom rules. Logic appears in the mathematics game as playing still another entirely different part: it furnishes the *rules of the game*. The only rule of moves is the following: If you have produced a formula $\mathfrak{b}$ and a formula $\mathfrak{b}\rightarrow\mathfrak{c}$ in which the same formula $\mathfrak{b}$ stands to the left of the sign →, you can put down the formula $\mathfrak{c}$. A mathematical proof consists in forming axioms according to the axiom rules and proceeding from them to new formulas by the repeated application of the syllogism rule just described. What is obtained in this manner are *provable* or rather *proved* formulas. One can judge from the looks of a complete genealogical tree of signs whether it is a formula or not. But one cannot judge from the looks of a complete formula whether it is provable. This is mainly caused by the fact that during its inductive construction a formula increases at every step, while in the syllogism two formulas $\mathfrak{b}$ and $\mathfrak{b}\rightarrow\mathfrak{c}$ combine to a new one $\mathfrak{c}$ which is shorter than the second premise, so that extension and contraction continually alternate in the proof game.

The axiom rules of equality, in which variables and the process of substitution already play a part, are the following:

$$\mathfrak{b}=\mathfrak{b}$$

$$(\mathfrak{b}=\mathfrak{c})\rightarrow(\mathfrak{A}(\mathfrak{b})=\mathfrak{A}(\mathfrak{c})).$$

They have an intermediate position between pure logic and mathematics. In the third place we have the purely arithmetical axioms with which we are well familiar and which relate to the notion of natural numbers, as for instance

$Z1$ (This is an individual definite axiom, not a rule for the formation of axioms.)

$Z\mathfrak{b}\rightarrow Z(\mathfrak{b}+1).$

Now comes the transfinite part of logic. With regard to "there exists," Σ_x, and "all," Π_x, we can for the time being only establish the following rules:

(*) $\mathfrak{A}(\mathfrak{b})\rightarrow\Sigma_x\mathfrak{A}(x)$ and $\Pi_x\mathfrak{A}(x)\rightarrow\mathfrak{A}(\mathfrak{b})$.

It is possible to infer existence from other assumptions, and it is possible to derive a particular application from a general assertion. But one cannot foresee at the moment how we can conclude in the opposite direction, how anything else can be inferred from the existence, or how a general assertion can follow from any other premise. I must describe how Hilbert extricates himself from this difficulty. (By the way, in our second axiom one can see very well how the German letters as tools of communication, so to speak, represent the hypothetically general, whereas the formal sign Π_x denotes the infinite logical product. This may help to clarify the distinction between the two ideas.) Let us take the property:

"x is honest" to be $\mathfrak{A}(x)$. If, in opposition to Brouwer, one appeals to the alternative that there must either be an honest man or that all men are dishonest, then an Aristides, who is the representative of honesty, can be found who is established to be honest if any man is honest. In the first case we choose for Aristides one of the honest men that exist, and in the second case an arbitrary one. But in order to be able to construct this Aristides, this representative, for every property, not only honesty, that is for every formula containing one single free variable x, we imagine that we have a divine automaton which accomplishes this task; when we insert the property $\mathfrak{A}$ into it, it produces the desired representative $\rho_x\mathfrak{A}$ which is sure to have the property $\mathfrak{A}$ if there is any individual of this kind. ρ_x is an integration sign. If we had an automaton like this at our disposal we would be free from the troubles caused by "all" and "any"; but of course the belief in its existence is pure nonsense. Mathematics, however, behaves as though it did exist. That can be expressed in an axiom rule, and the establishment of this rule is legitimate in formalized mathematics, provided its application does not lead to contradictions.

Thus we now add the following rules to (*):

$$\Sigma_x\mathfrak{A}(x)\rightarrow\mathfrak{A}(\rho_x\mathfrak{A}),\quad \mathfrak{A}(\rho_x(\backsim\mathfrak{A}))\rightarrow\Pi_x\mathfrak{A}(x).$$ [1]

Naturally they do not accomplish the same as the fictitious automaton, for it does not divulge what $\rho_x\mathfrak{A}$ is for a given formula $\mathfrak{A}$. Only in special cases a formula like $\rho_x\mathfrak{A}=1$ can result as final formula of a proof which starts from the axioms.

Now a few words concerning the second difficulty of transfinite character, the objectivation of properties. $\mathfrak{A}(x)$ being an arbitrarily given property, a formula involving the free variable x, the creation and existence of a new thing, the

---

[1] If $\mathfrak{A}(x)$ is "x is honest," then, $\mathfrak{A}(\rho_x(\backsim\mathfrak{A}))\rightarrow\Pi_x\,\mathfrak{A}(x)$ means "If the representative, $\rho_x(\backsim\mathfrak{A})$, of dishonesty is honest, then all men are honest."

"set" y, such that the proposition $\mathfrak{A}(x)$ is equivalent to $x \epsilon y$, must be expressed by a formal axiomatic rule:

$$\Sigma_y \, \Pi_x \, \{(x \epsilon y) = \mathfrak{A}(x)\}.$$

But it soon becomes evident that its application irrevocably leads to a contradiction. For classical analysis the limitation of the argument x to the domain of natural numbers is, however, sufficient; so that we establish the *transfinite rule for sets in the restricted form:*

$$\Sigma_y \, \Pi_x \, \{Zx \rightarrow ((x \epsilon y) = \mathfrak{A}(x))\}.$$

This rule is qualified to overcome the Russell types, just as the transfinite logical axioms guarantee the free manipulation of "all" and "any" prohibited by Brouwer for actual thinking.

I should now like to go into a short discussion of the consistency proof. The attempts to secure it have revealed the vicious circle character of mathematics to its full extent. Only after we have brought it to light completely can we succeed in finding the path that cuts through the circles and enables us to gain the insight that, despite the circles, no actual contradiction arises. This can be accomplished without trouble as long as the transfinite axioms are left aside. Under this restriction we are able to decide whether a normal formula is true or false by following its inductive construction. This on the one hand makes the indirect process of proof superfluous and on the other hand shows that it cannot lead to a contradiction. I indicate the rule by means of which we determine the truth-value T or F, true or false, of normal formulas. (A formula which does not involve the transfinite symbols Σ_x, Π_x, ρ_x, may be called a finite one). α) We always assign the value F to a basic formula, that is, to an individual constant or a variable. (It is of course immaterial whether we here decide in favor of T or F). β) A derived formula begins with a sign of integration or operation. In the first case—let us be magnanimous—we shall always assign to it the value T, likewise in the second case if the operation sign is not one of the following six: $\backsim \quad \vee \quad \& \quad \rightarrow \quad = \quad Z$. The evaluation of derived formulas beginning with one of these signs shall take place according to the following instructions:

(1)

$\mathfrak{b}$	$\backsim\mathfrak{b}$
T	F
F	T

(2)

$\mathfrak{b}$	$\mathfrak{c}$	$\mathfrak{b} \vee \mathfrak{c}$	$\mathfrak{b} \,\&\, \mathfrak{c}$	$\mathfrak{b} \rightarrow \mathfrak{c}$
T	T	T	T	T
T	F	T	F	F
F	T	T	F	T
F	F	F	F	T

(3) $\mathfrak{b} = \mathfrak{c}$ shall have the value T only when the two formulas $\mathfrak{b}$ and $\mathfrak{c}$ agree completely, and that again must be verified on their structure step by step. (4) A formula $Z\mathfrak{b}$ shall have the value T if no other signs except σ and 1 occur in the formula $\mathfrak{b}$. This rule constitutes our *rational evaluation.* It satisfies the following conditions:

(1) Every *finite* axiom has the value T;
(2) If $\mathfrak{b}$ and $\mathfrak{b} \rightarrow \mathfrak{c}$ have the value T, $\mathfrak{c}$ also has the value T.
(3) If $\mathfrak{b}$ has the value T, $\backsim\mathfrak{b}$ has the value F.

It follows that a proof which avails itself of finite axioms only consists entirely of formulas having the value T. If therefore a certain proof of such a kind ends with $\mathfrak{b}$, $\mathfrak{b}$ has the value T and $\backsim\mathfrak{b}$ the value F; it is thus impossible for any other proof of the same kind to end with $\backsim\mathfrak{b}$.

This argument may give you the impression that you are being mocked by a farce. But that is due to the fact that mathematics and mathematical proof are a farce as long as the transfinite axioms do not enter into the game. The formula $\mathfrak{b} \rightarrow \mathfrak{c}$ is always evaluated *after* $\mathfrak{b}$ and $\mathfrak{c}$ have been assigned truth-values. In the figure of the syllogism, this relation is inverted. We see from this that the syllogism is powerless without the transfinite axioms; the results it produces are much more easily reached by direct insight, that is, by the calculation of the value of the final formula as determined by its construction according to our evaluation rule. The syllogism would not save mathematics from being an immense tautology, but the transfinite is the vehicle which carries us beyond the domain of what is immediately conceivable.

Let us now add the transfinite logical axioms. It is convenient to use our four axioms separately in building up mathematics, but for the consistency proof it is more suitable to replace them by a single one in which their entire force is concentrated:

$$(\infty) \qquad \mathfrak{A}(\mathfrak{b}) \rightarrow \mathfrak{A}(\rho_x \mathfrak{A})$$

and to use only the transfinite symbol ρ_x instead of our three, ρ_x, Σ_x, Π_x. It is surely out of the question to arrange the descriptive valuation rule in such a manner that we can be assured that all possible axioms formed according to the rule (∞) receive the value T; this signifies that the insight into true and false has here come to an end. However, less than this is sufficient for the consistency proof. We assume that two proofs are concretely given which lead to a formula $\mathfrak{b}$ and its negation $\smallsmile\mathfrak{b}$. Let the axioms, given explicitly and containing nothing undetermined, which are used in these two proofs, be briefly designated by

$$(\dagger) \qquad \mathfrak{s}: \quad \mathfrak{s}_1, \mathfrak{s}_2, \cdots, \mathfrak{s}_n.$$

In order to convince oneself that the situation assumed above cannot possibly come to pass it suffices to find an evaluation which is "correct within $\mathfrak{s}$." I mean by this that the rule of evaluation shall satisfy the conditions (2) and (3) as before, but instead of (1) we shall only require that *the* *few axioms of our set* $\mathfrak{S}$ *shall have the truth value T.* This "artificial evaluation" will naturally depend on that system $\mathfrak{S}$. Stated more precisely, our problem is *to construct by means of a definite universal method an evaluation for any set* $\mathfrak{S}$ *of arbitrary, but explicitly given, axioms* (†), *which is correct within* $\mathfrak{S}$.

I here consider the simplest case of all: wherever the transfinite symbol ρ occurs in those particular cases of the rule (∞) which are contained in our given set $\mathfrak{S}$ it shall contain the same index x and it shall always govern the *same* definite *finite* formula $\mathfrak{E}$ ($\mathfrak{E}$ of course contains only x as a free variable). If you like you may call this property $\mathfrak{E}$ "honesty." Let $\mathfrak{b}$ be a given finite normal formula; we try whether $\mathfrak{b}$ can be used as representative of $\mathfrak{E}$—as "Aristides;" i.e., we replace $\rho_x\mathfrak{E}$, wherever it occurs, by $\mathfrak{b}$. By this process there arises from an arbitrary formula $\mathfrak{c}$ the "reduced formula" $\hat{\mathfrak{c}}$. We call this "the reduction $\mathfrak{b}$" and give to $\mathfrak{c}$ the value T or F according as the reduced formula $\hat{\mathfrak{c}}$ is given the value T or F by the rational evaluation. The only point which is to be tested is whether the special cases of the transfinite axiom rule (∞) which occur in the system $\mathfrak{S}$ are given the value T. They contain, by assumption, ρ only in the same combination $\rho_x\mathfrak{E}$; let them be the following

(A) $\mathfrak{E}(\mathfrak{b}_i)\rightarrow\mathfrak{E}(\rho_x\mathfrak{E})$ $(i=1, 2, \cdots, h)$

$\mathfrak{b}_1, \mathfrak{b}_2, \cdots, \mathfrak{b}_h$ need not be finite, but can themselves contain the transfinite symbol ρ in the combination $\rho_x\mathfrak{E}$.

We try for luck the reduction 1 and the corresponding evaluation; the reduction symbol ^ may now refer to this case. The original mathematical object 1, Adam, shall be our Aristides. The formulas (A) are changed into (Â) by this reduction: $\mathfrak{E}(\hat{\mathfrak{b}}_i)\rightarrow\mathfrak{E}(1)$. We distinguish between several cases. The most favorable is that in which Adam stands the test, in which the finite formula $\mathfrak{E}(1)$ is true, i.e., is given the truth-value T in the rational evaluation. If, however, $\mathfrak{C}(1)$ is false, if Adam is a scoundrel, our first unhappy choice does no harm provided none of the finite formulas

$$\mathfrak{C}(\hat{\mathfrak{b}}_1),\quad \mathfrak{C}(\hat{\mathfrak{b}}_2),\ \cdots,\ \mathfrak{C}(\hat{\mathfrak{b}}_h) \tag{B}$$

have the value T: both sides of these relations have then the value F, and then themselves have therefore the value T. It may have been absurd to choose 1 as the representative of $\mathfrak{C}$ from the standpoint of absolute truth; Adam is dishonest and in spite of that there may be honest men. But it need not disturb us if none of the few definite people $\hat{\mathfrak{b}}_1, \hat{\mathfrak{b}}_2, \cdots, \hat{\mathfrak{b}}_h$ with whom we are now dealing in these axioms, are not scoundrels. The only case in which our evaluation 1 does not lead to our goal is that in which $\mathfrak{C}(1)$ is false but one of the finite formulas (B), e.g. $\mathfrak{C}(\hat{\mathfrak{b}}_1)$ is true. In this case, however, we possess a legitimate representative of the property $\mathfrak{C}$ in $\hat{\mathfrak{b}}_1$. Thereupon we reject the first process of reduction and instead of it perform a second one $\hat{\mathfrak{b}}_1$, i.e., we replace $\rho_x\mathfrak{C}$ wherever it occurs by $\hat{\mathfrak{b}}_1$. So if the first reduction chosen at random fails, its very failure gives us automatically another reduction which leads us to our goal.

But we only stand at the very beginning of the complications that can arise. ρ can combine with different $\mathfrak{A}$'s to $\rho_x\mathfrak{A}$, the $\mathfrak{A}$'s can again contain this transfinite symbol, so that the ρ's are heaped onto one another. When a certain reduction fails, the failure does automatically furnish a new reduction which is successful *there* where the original one failed. But in return for that it will in general go wrong with regard to those axioms of the set $\mathfrak{s}$, for which the first one worked; so that on the whole we are not at all sure that this auto-correction really produces an improvement. The point to be proved is that one does not turn about in a circle, but that after a number of steps of successive corrections which can be indicated at the outset, a reduction is arrived at which does not fail anywhere within the given finite set of axioms $\mathfrak{s}$. The successive corrections can be considered as steps in a certain combinatorial game; and we assert that the game comes necessarily to an end after a finite number of steps, however we play, our freedom being restricted only by the rules of the game. J. v. Neumann has proved this theorem of finiteness. We are here dealing with a concrete mathematical problem which is not trivial, but at the same time is solvable, and I cannot imagine that any mathematician can find the courage to elude its honest solution by means of a metaphysical dogma.

Thus Mr. v. Neumann succeeded in proving that one does not encounter any contradictions in a "restricted analysis," where one handles the sequence of natural numbers as though it were a closed set of objects existing in themselves. The justification of the same procedure for the continuum of real numbers would require the proof that the restricted transfinite axiom concerning sets does not introduce a contradiction into the system of axioms. This problem is surely of a much deeper nature. As I hear, Mr. Ackermann, a pupil of Hilbert's, has reduced it to a form similar to that which Neumann obtained for the restricted analysis. But he has not yet obtained, and it is perhaps even permitted to doubt whether his combinatorial game, which is of course of a rather complicated character, really possesses, the requisite property of finiteness. Thus the situation remains serious but not entirely hopeless for classical analysis.

II

THE SPHERICAL SYMMETRY OF ATOMS

An atom or an ion, whose nucleus is considered as a fixed center of force O, possesses two kinds of symmetry properties: (1) the laws governing it are spherically symmetric, i.e., invariant under an arbitrary rotation about O; (2) it is invariant under permutation of its f electrons. The first type of symmetry is described by the continuous group[1] γ of orthogonal transformations of the spacial coordinates x, y, z, and the second by the finite symmetric group of order $f!$ which consists of all permutations of f things. This accounts for the importance of these two groups in physics. It is my intention to give a general idea of what the first kind of symmetry means for quantum physics. But before taking up these applications of group theory I find it desirable to describe the general foundations of quantum mechanics as briefly as possible.

Let the physical system P under consideration be, as I said before, an atom or an ion consisting of a fixed nucleus in O and f electrons revolving about it. I employ Cartesian coordinates x, y, z with origin in O as the coordinates of the three-dimensional physical space. For the mathematical representation of the state of P we need (in addition to the actual space) an abstract system space which has in general an infinity of dimensions: it is to be a *unitary vector space*. What does that mean? In an n-dimensional Euclidean vector space each vector $\mathfrak{x}$ is described with respect to a Cartesian or normal coordinate system by its n components $x_1, x_2, \cdots, x_n$. The square of the length of $\mathfrak{x}$ is

---

[1] For a more extensive treatment of groups see H. Weyl, *Gruppentheorie und Quantenmechanik*, Leipzig, 1928. This book is hereafter referred to as GQ.

$$\mathfrak{x}^2 = x_1^2 + x_2^2 + \cdots + x_n^2. \tag{1}$$

All Cartesian coordinate systems are equally admissible; they are transformed into each other by *orthogonal transformations*, i.e., linear homogeneous transformations which leave the quadratic form (1) invariant. But when the components x_i are allowed to take arbitrary complex values instead of real ones, as is the case in quantum physics, the form (1) loses its positive definite character and must be replaced by

$$\mathfrak{x}^2 = \bar{x}_1 x_1 + \bar{x}_2 x_2 + \cdots + \bar{x}_n x_n \tag{2}$$

where $\bar{x}_i$ is the conjugate complex of x_i. We then speak of unitary instead of Euclidean geometry; linear transformations which leave $\bar{x}_1 x_1 + \cdots + \bar{x}_n x_n$ invariant are called unitary transformations. The scalar product $(\mathfrak{x}\ \mathfrak{y})$ of two vectors $\mathfrak{x}$ and $\mathfrak{y}$ in unitary geometry is defined by

$$(\mathfrak{x}\ \mathfrak{y}) = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n.$$

The first general principle in the construction of the mathematical skeleton[1] of quantum theory is this:

I. *There is associated with the physical system P a unitary system space* $\mathfrak{R}$; *every state of the system is represented by a vector in* $\mathfrak{R}$. In Schrödinger's wave theory of a single particle this space $\mathfrak{R}$ consists of all functions $\psi\ (x\ y\ z)$ of the coordinates $x\ y\ z$ of the particle for which the integral of $\bar{\psi}\psi$ has a finite value. They form a vector space (of an infinity of dimensions) because these functions allow addition and multiplication with arbitrary complex numbers. The integral just mentioned replaces the sum (2). The undulatory character of physical phenomena, the fact that they show *interference*, i.e., linear superposition with arbitrary shifts of phase, means that we are dealing with vectors in a complex vector space which are capable of addition and multiplication by complex numbers. It is unitary because there exists something like *intensity.*

---

[1] See GQ, §14.

A linear transformation

$$x_i' = \Sigma_k \, l_{ik} x_k$$

can be interpreted as an *operator* L which carries the arbitrary vector $\mathfrak{x} = (x_i)$ over into $\mathfrak{x}' = (x_i')$; the linear character of such an operator is described by the functional equations

$$(\mathfrak{x} + \mathfrak{y})' = \mathfrak{x}' + \mathfrak{y}'; \quad (a \cdot \mathfrak{x})' = a \cdot \mathfrak{x}'$$

$L = \| l_{ik} \|$ is the coefficient matrix of L in the chosen normal coordinate system. If we go over to a new normal coordinate system by means of the unitary transformation U the same operator L is expressed by another matrix $L' = \| l'_{ik} \|$ instead of L, namely

$$L' = ULU^{-1}.$$

L is itself unitary if it leaves the square of the length of an arbitrary vector invariant: $\mathfrak{x}'^2 = \mathfrak{x}^2$. L is an Hermitian operator if the scalar product $(\mathfrak{x} \cdot L\mathfrak{y})$ is conjugate complex to the scalar procuct $(\mathfrak{y} \cdot L\mathfrak{x})$; the matrix L then fulfills the condition of symmetry $l_{ki} = \bar{l}_{ik}$. The scalar product

$$(\mathfrak{x} \cdot L\mathfrak{x}) = \Sigma l_{ik} \bar{x}_i x_k$$

is called an Hermitian form.[1]

The main theorem on Hermitian forms is the theorem concerning the existence of principal axes, which is so familiar to you in the two- and three-dimensional Euclidean geometry, where ellipses and ellipsoids appear in place of Hermitian forms. One can always find a normal coordinate system $\mathfrak{e}_i$ in which the given Hermitian form L reduces to

$$\Sigma_i \, l_i \, \bar{x}_i \, x_i$$

or in terms of which the Hermitian operator is expressed by the simple equation $x_i' = l_i x_i$. $\mathfrak{e}_i$ are the characteristic vectors

---

[1] GQ, p. 17.

("Eigenvektoren" in German) and l_i the corresponding characteristic numbers ("Eigenwerte").

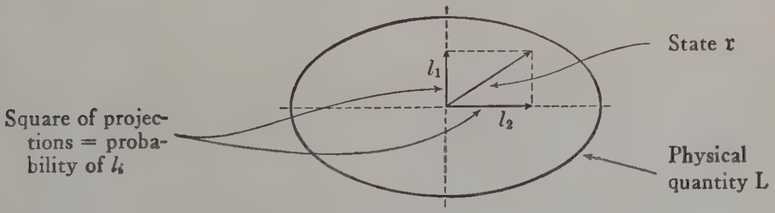
I can express the second and third general principles of quantum theory in terms of these notions. The general question of quantum theory is this: What are the possible values of a given physical quantity L, and what is the probability that it will assume these values when the system is in a given state $\mathfrak{x}$? Before we can answer this question we must know how the state of our physical system and how any quantity L of P can be described mathematically. Principle I takes care of the description of the state $\mathfrak{x}$, and Principle II asserts that:

II. *Each physical quantity L of P is represented by a certain Hermitian operator L in system space* $\mathfrak{R}$.[1]

And on the basis of this mathematical scheme the answer to the general physical question is:

III. *The possible values of a physical quantity L are the characteristic values* l_i *of the Hermitian form L which represents it; the relative probability that it will assume the value* l_i *in the state* $\mathfrak{x}$ *is given by the number* $\bar{x}_i x_i$, *i.e., the square of the absolute value of the projection of* $\mathfrak{x}$ *along the i-th principal axis.*[2]

The situation is the following—*cum grano salis!* (The diagram has been drawn in two instead of an infinite number of dimensions.)



The quantity under consideration will only assume a value with certainty if $\mathfrak{x}$ falls along one of the principal axes.

---

[1] GQ, §14. [2] GQ, §5.

In Schrödinger's theory[1] of a single particle the coordinate x is represented by the linear Hermitian operator which carries ψ into $x\psi$ (multiplication by x). This operator $\psi \rightarrow x\psi$ is already referred to principal axes; Schrödinger employs that coordinate system in system space in which the operators which represent the coordinates of the electron are referred to principal axes. The probability that the coordinates $x\ y\ z$ of the electron are within a volume V is given by the integral of $\bar{\psi}\psi$ over V, in agreement with our general probability principle III.

The components p_x, p_y, p_z of momentum are represented by the Hermitian operators

$$p_x : \psi \rightarrow \frac{h}{2\pi i}\,\frac{d}{dx}\psi; \cdots$$

in Schrödinger's theory. h is Planck's constant; I shall henceforth employ units which are so chosen that $h/2\pi$ becomes equal to 1.

The fourth principle is the *dynamical law*[2] which asserts that:

IV. *The state* $\mathfrak{x}$ *of the system goes over into a new state* $\mathfrak{x}+d\mathfrak{x}$ *after lapse of time* dt *which arises from* $\mathfrak{x}$ *by an infinitesimal unitary operator*

$$d\mathfrak{x} = \frac{dt}{i}\,H\mathfrak{x} \quad \text{or} \quad \frac{d\mathfrak{x}}{dt} = \frac{1}{i}\,H\mathfrak{x}.$$

If I write an arbitrary infinitesimal unitary operator in this form it can be readily shown that H is an Hermitian operator. The quantity represented by H is called *energy*.

I now come back to my proper subject. In addition to the *real* process which carries the state $\mathfrak{x}$ of our physical system

---

[1] GQ, §10.

[2] GQ, §15.

P at time t over into the new state $\mathfrak{x}'$ at a later time t', I consider a *virtual* process: I rotate P about the origin O by the rotation s; by this process the arbitrary state $\mathfrak{x}$ of P goes over into a new state $\mathfrak{x}'$. Since everything which is physically significant must remain unchanged and since the system space of the vector $\mathfrak{x}$ is unitary, the transition $\mathfrak{x}\rightarrow\mathfrak{x}'$ must be a linear unitary operator $U(s)$, dependent on s, in the system space. I say briefly that $U(s)$ is *induced* in the system space by s. All such rotations s form the group of rotations of orthogonal transformations in our 3-dimensional Euclidean space. The correspondence $U : s\rightarrow U(s)$ must obey the following law

$$U(t)U(s)=U(ts):$$

To the composition of rotations s, t corresponds the composition of the induced operations U in system space. The mathematician calls such a correspondence a *representation*[1] of the rotation group.

The transition $t\rightarrow t'$ occurring in real physical time can be generated by the infinitesimal transition $t\rightarrow t+dt$, the dynamical law giving the induced infinitesimal unitary operator in system space: $d\mathfrak{x}=\frac{dt}{i}H\mathfrak{x}$. In the same way we may generate the group of rotations by the infinitesimal rotations about the x-, y-, and z-axes. The infinitesimal rotation about the z-axis is given by

$$\delta x=-y,\quad \delta y=x,\quad \delta z=0. \tag{3}$$

I have purposely written no infinitesimal constant factor e on the right hand side of these equations; you may interpret the differential symbol δ as differentiation $\frac{d}{d\tau}$ with respect to a fictitious or virtual time τ. Let the infinitesimal unitary

[1] GQ, §24.

operator induced in system space by this infinitesimal rotation be

$$\delta\mathfrak{x} = -iM_z\mathfrak{x}.$$

M_x, M_y, M_z are then Hermitian operators; I call them, or rather the physical quantities represented by them the components of angular momentum M. Angular momentum has therefore the same significance for the virtual alteration of $\mathfrak{x}$ by rotation as the energy has for the real alteration of the system P in time.

It can be easily deduced from the dynamical law that a physical quantity A is constant in time when its representing Hermitian operator commutes with $H : HA = AH$. That it is constant here means that the probabilities with which the quantity assumes its different values do not change in time. Of course H commutes with H itself, and therefore satisfies the law of conservation. But in the same way we see that A is a scalar quantity, i.e., is unchanged by rotation in space, if and only if it commutes with M, i.e., with the unitary operator induced in system space by the rotations in physical space. Now the energy H is certainly a scalar quantity, and consequently M commutes with H. But in accordance with what I have said above this means that M is constant in time; angular momentum satisfies the law of conservation.

I shall illustrate these very general and abstract considerations by the simplest case which fits into the general scheme: Schrödinger's scalar wave theory for a single particle. Let the rotation s carry the point $x\ y\ z$ over into $x'\ y'\ z'$; into what state ψ' does it carry the state ψ of our particle described by the Schrödinger function $\psi\ (x\ y\ z)$? Obviously this new state is defined by the equation

$$\psi'(x'y'z') = \psi(xyz). \tag{4}$$

If s is an infinitesimal rotation, the increments

$$\psi'(xyz)-\psi(xyz)=\delta\psi \text{ and } \delta x=x'-x$$

are all infinitesimals; our equation (4) then yields

$$\delta\psi+\left(\frac{\partial\psi}{\partial x}\,\delta x+\cdots\right)=0.$$

For the infinitesimal rotation (3) about the z-axis we get

$$\delta\psi=-\left(x\frac{\partial\psi}{\partial y}-y\frac{\partial\psi}{\partial x}\right)$$

and consequently M_z is the operator

$$L_z=\frac{1}{i}\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right) \tag{5}$$

and this is in complete agreement with the expression for angular momentum in classical dynamics:

$$L_z=xp_y-yp_x \tag{5*}$$

I believe that the validity of the conservation law and this agreement with classical mechanics affords sufficient justification for our definition of angular momentum.

A unitary vector space $\mathfrak{R}$ can be decomposed[1] into mutually perpendicular linear subspaces $\mathfrak{R}=\mathfrak{R}_1+\mathfrak{R}_2+\cdots$. That is to say, any vector $\mathfrak{x}$ in $\mathfrak{R}$ can be written in one and only one way as a sum $\mathfrak{x}_1+\mathfrak{x}_2+\cdots$ the individual summands of which lie in $\mathfrak{R}_1$, $\mathfrak{R}_2$, $\cdots$ respectively. For example, three-dimensional space can be decomposed into a plane and the line perpendicular to it. If we are given a linear unitary transformation of space, or a group of such transformations, we wish to carry out the decomposition in such a way that each of the subspaces $\mathfrak{R}_1$, $\mathfrak{R}_2$, $\cdots$ remains invariant under the given transformations; we further wish to carry it as far as possible under this condition. We then say: $\mathfrak{R}$ is decomposed into irreducible invariant sub-spaces with respect to the group under consideration. If, for example, we have but a single infinitesimal unitary transformation the theorem on principal axes tells us that we can decompose into one-dimensional invariant sub-spaces—along the principal axes. The decomposition into irreducible invariant sub-spaces consequently corresponds to the separation of the various values which are possible for a physical quantity.

---

[1] GQ, pp. 10-11.

In particular, we consider the decomposition into invariant sub-spaces $\mathfrak{R}$ relative to the group of transformations which are induced in system space $\mathfrak{R}$ by the group of rotations in actual space. The rotation group then induces a definite representation Γ_j in each sub-space $\mathfrak{R}_j$. The operator representing moment of momentum is correspondingly separated into partial operators, each of which operates on one of the spaces $\mathfrak{R}_j$. At this point we use to advantage that discipline which the mathematicians call Topology or Analysis Situs. The rotations build a closed manifold, like the points on the surface of a sphere, and this has as consequence that there exists but a discrete set of different irreducible representations Γ_j of the group of rotations, which can be distinguished from one another by an index j which runs through the values $j=0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$. The space $\mathfrak{R}_j$ in which the representation Γ_j with index j is induced has $2j+1$ dimensions. Consequently we know the moment of momentum M in $\mathfrak{R}_j$ independently of the dynamical structure of the physical system under consideration; its components are the operators which correspond to the infinitesimal rotations in the representation Γ_j. Computing the square of the absolute value M^2 we find that in $\mathfrak{R}_j$ this operator is simply multiplication by the constant $j(j+1)$. We thus come to the conclusion that the total system space $\mathfrak{R}$ can be decomposed into subspaces $\mathfrak{R}_j$ in such a way that

(1) $\mathfrak{R}_j$ has $2j+1$ dimensions;

(2) the energy in $\mathfrak{R}_j$ has a definite value E_j; and

(3) the square of the moment of momentum M^2 has the definite value $j(j+1)$.
j is called the inner quantum number of an atomic state represented by a vector lying in $\mathfrak{R}_j$. The infinitesimal displacement in time induces in system space the operator representing energy. Decomposition into invariant subspaces with respect to the group of actual displacements in time leads, therefore, to the principal axes of the energy operator. But since these displacements constitute an open group the various possible one-dimensional representations are not restricted *a priori* to a definite discrete set. This is the reason that we can give *a priori* definite discrete values of which the moment of momentum is capable, but not so in the case of the energy.

I must now mention as the last, the V-th quantum theoretic principle, that one which tells us how to effect the *composition of two physical systems*[1] into a single one. If (x_i) be an arbitrary vector in an m-dimensional space $\mathfrak{R}$, and (y_k) one such in an n-dimensional space $\mathfrak{R}'$, we can consider the $m \cdot n$ numbers $x_i y_k$ as the components of a vector $z = \mathfrak{x} \times \mathfrak{y}$ in an $(m \cdot n)$-dimensional space which I designate as the cross-product $\mathfrak{R} \times \mathfrak{R}'$. If the x_i be subjected to a linear transformation A and the y_k to B, then the $m \cdot n$ linearly independent quantities $x_i y_k$ undergo a certain linear transformation $A \times B$ induced in the product space $\mathfrak{R} \times \mathfrak{R}'$ by A in $\mathfrak{R}$ and B in $\mathfrak{R}'$. The rule of composition is: if $\mathfrak{R}$, $\mathfrak{R}'$ be the system spaces of the physical systems P, P' respectively, then $\mathfrak{R} \times \mathfrak{R}'$ is the system space of the total system composed of both P and P'. An operator A which represents a physical quantity of P in $\mathfrak{R}$ must be replaced by $A \times 1$ in order to represent the same quantity in the total system space $\mathfrak{R} \times \mathfrak{R}'$; $\times 1$ is here the identity in $\mathfrak{R}'$. This factor $\times 1$

---

[1] GQ, §17.

leaves all relations between the operators A undisturbed. Similarly, an operator B which represents a quantity of P' in $\mathfrak{R}'$, must be replaced by $1\times B$. It is thus possible to represent the quantities of P and P' by operators in the *same* system space $\mathfrak{R}\times\mathfrak{R}'$ without affecting the relations which exist between the quantities of either system. This is the situation intended when we say that the total physical system consists of two *kinematically independent parts.*

What influence has multiplicative composition on the infinitesimal operators? The equation

$$d(x_i y_k) = dx_i \cdot y_k + x_i \cdot dy_k$$

shows that the infinitesimal transformations

$$dx = Ax, \quad dy = By$$

in $\mathfrak{R}$ and $\mathfrak{R}'$ induce the infinitesimal transformation

$$(A\times 1)+(1\times B)$$

in the product space. In this sense infinitesimal operators behave additively on multiplicative composition.

Let the group γ of rotations induce representations

$$\mathfrak{U} : s\rightarrow U(s) \text{ in } \mathfrak{R}, \quad \mathfrak{U}' : s\rightarrow U'(s) \text{ in } \mathfrak{R}';$$

it therefore induces the representation

$$\mathfrak{U}'\times\mathfrak{U}' : s\rightarrow U(s)\times U'(s)$$

in product space $\mathfrak{R}\times\mathfrak{R}'$. Applying this to infinitesimal rotations and using the formula just derived for such we obtain the theorem: The moment of momentum of a physical system which consists of two kinematically independent parts is equal to the sum of the moments of momentum of the two partial systems. One might think that the same argument would hold for the operators in system space which give the infinitesimal displacement of the state vector in time, i.e., that the energy of the whole is equal to the sum of the energies of the parts. But there is an essential distinction between these two cases; the rotation in space is a virtual process, whereas the displacement in time describes what actually happens. And this has as consequence that the law of the addition of energy only holds when the partial systems are dynamically, as well as kinematically, independent, i.e., when there is no interaction between them. The same theorem for angular momentum is not bound to this condition.

If the one part has an angular momentum whose magnitude is j and the other part j', then the angular momentum of the whole is, on classical mechanics, capable of all values between the limits $j+j'$ and $|j-j'|$. The various possibilities are conditioned by the various angles which the two moments, vectors of lengths j and j', can make with each other. What is the analogous problem in quantum mechanics? We must build the cross-product of the two irreducible representations Γ_j, Γ_j' and decompose it into its irreducible constituents. The problem is solved by the formula[1]

$$\Gamma_j \times \Gamma_j' = \Sigma_\iota \, \Gamma_\iota$$

where ι takes on each of the values $j+j'$, $j+j'-1, \cdots,$ $|j-j'|$ exactly once.

In order to follow the construction of an atom from electrons we will find it convenient to make use of a *vector model,* into which the angular momenta of the single electrons are introduced and added together in order to obtain the total. The mathematical interpretation given this model by quantum mechanics is characterized by the two circumstances:

(1) The determination of the various numerical possibilities is to be interpreted as decomposition into invariant irreducible sub-spaces;

(2) The addition of vectors has its mathematical counterpart in the multiplicative composition of the representations induced in these sub-spaces.

---

[1] GQ, §30.

This interpretation accounts for the following deviations from classical mechanics:

1) The inner quantum number j is restricted to the discrete values $0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$;

2) The square of the absolute value of the angular momentum is $j(j+1)$ instead of j^2;

3) The inner quantum number obtained on composition is not capable of all values between $j+j'$ and $|j-j'|$, but only of those among them which differ from $j+j'$ by an integer.

In Schrödinger's wave theory the inner quantum number is only capable of integral values. This is, however, in disagreement with the observations on alkali spectra, which show a doublet with inner quantum numbers $j=l\pm\frac{1}{2}$ instead of a simple term with inner quantum number l. But if the wave function ψ has two components ψ_1, ψ_2 which depend on the coordinate system in such a way that under the influence of a rotation s they undergo that transformation which corresponds to s in the representation $\Gamma_{\frac{1}{2}}$, these facts are readily explained. The situation can be described by considering the single electron to consist in abstract of two kinematically independent parts, the electron translation and the electron spin.[1] The arguments (x, y, z) in ψ characterize the components in the system space associated with the translation and the indices 1, 2 the components in the two-dimensional system space associated with the spin. The angular momentum of translation is given by the classical formula (5); its inner quantum number is always integral and is called the azimuthal quantum number of the electron. The inner quantum number of the spin is $\frac{1}{2}$. If we distinguish in this way between the translatory angular momentum and the spin of each electron we arrive naturally at the Hund vector model of the atom which solved the problem of atomic line spectra with such a definite success. But this model is not to be taken literally; we obtain it in a new interpretation, in which the deviations 1)—3) from classical mechanics are explained from a unified viewpoint and not by hypotheses added *ad hoc*. This is the service rendered by the new quantum theory.

---

[1] GQ, §37.

I have taken pains to show clearly the simple correspondence between the language of classical mechanics and that of quantum theory. There can be no doubt that to the same extent as the use of vectors, capable of all magnitudes and directions, and their additive composition were appropriate mathematical instruments for classical mechanics, so are decomposition into invariant sub-spaces relative to a given group and multiplicative composition of its representations the basic mathematical operations for the description of the same physical phenomena in the new mechanics. But far be it from me to exaggerate the importance of such a mathematical language. It is the same as with ordinary language; it is true that it can be of considerable assistance to thought, which should strictly be alone directed on the subject, in that it assists in fixing partial knowledge already won for further use. But it is also true that it entails the danger of carrying the spirit away from the subject proper into the void; poets and philosophers have occasionally made a virtue of such a predicament. But, after all, the clothing of physical relationships in a mathematical symbolism adapted for other purposes is not to be tolerated for long; hence it would be ungrateful of physics to refuse the valet service performed by mathematics in divesting it of an outworn and shabby garment. The foregoing is what I wish to present here for your consideration as an illustration of the applications of group theory to the new quantum mechanics.

III

GRAVITATION AND THE ELECTRON[1]

In order to describe physical quantities by numerical data we must refer them to a Cartesian system of axes in space which consists of three mutually perpendicular vectors e of length 1. Transition to another allowable set of axes is accomplished by an orthogonal transformation or rotation

$$e'_\alpha = \Sigma_\beta \, O_{\alpha\beta} \, e_\beta. \tag{1}$$

Vectors and tensors are quantities which are determined relative to an axis system by a set of numbers, their components, in such a way that these components are transformed in a definite way on transition to another axis system. More precisely, the linear transformation of the components is related to the arbitrary rotation (1) by a certain law in such a way that to composition of two rotations corresponds the composition of the two associated transformations. Such a correspondence between the elements of a group—here the group of rotations—and linear transformations is called a representation of the group. *Each kind of quantity is characterized by a definite representation of the rotation group.* However the familiar vectors and tensors are not the only quantities of this kind; a possibility which had not cropped up previously in physics is necessary for the description of the electron spin in wave mechanics.

A rotation is a transformation of the unit sphere into itself. The sphere can be projected stereographically onto its equatorial plane, which is to be considered as described by the complex variable $\zeta = x + iy$. We then replace the one complex coordinate ζ on the sphere by two homogeneous coordinates ψ_1, ψ_2 by writing $\zeta = \psi_2/\psi_1$ (in order to include in the representation the center of projection on the sphere). The formulas for the stereographic projection are then (x_α being the Cartesian coordinates in space)

---

[1] For similar treatments see H. Weyl, *Gravitation and the Electron*, Proc. Nat. Acad. Sciences, V, 15, No. 4, pp. 323–34, 1929, and *Elektron und Gravitation I*, Zeit. für Phy., 56 Bd., 5 & 6 Hft., p. 330, 1929. Full references are included in these papers.

$$\begin{array}{lccccccc} & x_1 & : & x_2 & : & x_3 & : & 1 \quad = \\ (2) & \psi_2\bar{\psi}_1 + \psi_1\bar{\psi}_2 & : & \frac{1}{i}(\psi_2\bar{\psi}_1 - \psi_1\bar{\psi}_2) & : & \psi_1\bar{\psi}_1 - \psi_2\bar{\psi}_2 & \cdot & \psi_1\bar{\psi}_1 + \psi_2\bar{\psi}_2 \end{array}$$

Each rotation s of the sphere can now be represented by a unitary transformation of the coordinates ψ_1, ψ_2, i.e., by a linear transformation (s) which leaves $\psi_1\bar{\psi}_1 + \psi_2\bar{\psi}_2$ invariant (Cayley and Helmholtz). Of course this transformation (s) is only definite in the homogeneous sense, so that transformations which are obtained from each other by multiplying all coefficients by a number $e^{i\lambda}$ of absolute value 1 are to be considered as the same. We can avoid this arbitrariness by a normalization, requiring that the determinant of the transformation shall be $+1$; it is still double-valued, however, since multiplication of all coefficients by -1 does not disturb the normalization. (2) are Hermitean forms in ψ_1, ψ_2 with the coefficient matrices

$$(3) \qquad S_1 = \left\|\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}\right\|, \quad S_2 = \left\|\begin{matrix} 0 & -i \\ i & 0 \end{matrix}\right\|, \quad S_3 = \left\|\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}\right\|.$$

Denoting by ψ the matrix with the one *column* ψ_1, ψ_2 and by $\bar{\psi}$ the matrix with one *row* consisting of the conjugate complex quantities $\bar{\psi}_1$, $\bar{\psi}_2$, the expressions (2) for the rectangular components become

$$x_\alpha = \bar{\psi} S_\alpha \psi.$$

The law of transformation of the ψ is consequently determined by the requirement that the three quantities $\bar{\psi} S_\alpha \psi$

transform like the components of a vector. A quantity having two components with this law of transformation describes the wave field of an electron in Pauli's theory of the electron spin; both components of ψ are here of course functions of position (and time).

It follows immediately from the above law that

$$\Sigma_\alpha S_\alpha \frac{\partial}{\partial x_\alpha} = \nabla$$

is a differential operator independent of coordinate system which transforms ψ into quantity $\nabla\psi$ of the same kind. Dirac remarked that the reiterated operator $\nabla\nabla$ is the Poisson operator

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

He therefore replaced the de Broglie-Schrödinger wave equation of second order for the scalar quantity ψ, which cannot take account of the spin, by two first order equations in the two components ψ_1, ψ_2 in which this differential operator ∇ plays the same rôle as the Poisson operator in the scalar theory, and was in this way able to give a most satisfactory explanation of the anomolous Zeeman effect in all its details.

But he did still more. We must really operate in the four-dimensional world—instead of in three-dimensional space—which has, in accordance with the special theory of relativity a geometry (Minkowski's geometry) analogous to that of three-dimensional Euclidean space. A Cartesian system of axes there consists of three real space-like vectors $e(1)$, $e(2)$, $e(3)$ and a pure imaginary vector $e(0)$; we expressly demand that the real time-like vector $e(0)/i$ point toward the future. Unfortunately the quantity ψ has then four components ψ_1^+, ψ_2^+; ψ_1^-, ψ_2^- instead of two. $S(\alpha)$ are now linear transformations of the four components under which the two +-components and the two −-components transform among themselves; for the +-components it equals the S_α given above and differs only in sign for the −-components. To them must be added

$$S(0) = \begin{Vmatrix} i & 0 \\ 0 & i \end{Vmatrix}$$

The law of transformation of the components of ψ is again to be described as one under which the four quantities $\bar{\psi} S(\alpha) \psi$ transform like the components of a four-vector under transition from one Cartesian coordinate system to another.

The Dirac wave equations arise from an *action principle* which is additively composed of the two integral invariants:

$$(4) \qquad \frac{h}{i} \int \bar{\psi} S(\alpha) \frac{\partial \psi}{\partial x_\alpha} dx \qquad (dx = dx_0 dx_1 dx_2 dx_3)$$

(α summed over 0, 1, 2, 3), and

$$(5) \quad cm \int (\psi_1^+ \bar{\psi}_1^- + \psi_2^+ \bar{\psi}_2^- + \psi_1^- \bar{\psi}_1^+ + \psi_2^- \bar{\psi}_2^+) dx = cm \int (\bar{\psi}^- \psi^+ + \bar{\psi}^+ \psi^-) dx$$

m being the mass of the electron. The components of ψ are the quantities which are to be varied. This holds as long as there is no electromagnetic field present; if there be such, the operator $\frac{\partial}{\partial x_\alpha}$ acting on ψ must be replaced by $\frac{\partial}{\partial x_\alpha} + \frac{ie}{h} \varphi_\alpha$ ($-e$ the charge of the electron, $\frac{h}{2\pi}$ the quantum of action and φ the components of the electromagnetic potential). In addition to Dirac's equation we have Maxwell's equations, in which the four-vector of charge-current density is given by

$$(6) \qquad \rho_\alpha = -e \bar{\psi} S(\alpha) \psi$$

in particular the charge density is

$$(7) \qquad -e(\psi_1^+ \bar{\psi}_1^+ + \psi_2^+ \bar{\psi}_2^+ + \psi_1^- \bar{\psi}_1^- + \psi_2^- \bar{\psi}_2^-)$$

This corresponds to the circumstances that the Maxwellian action of the electromagnetic field is to be added to (4) and (5) as a third constituent and the φ_α also varied.

These field equations differ strikingly from those of classical physics in that *they contain the fundamental atomic constants* e, h, m. Denoting $\frac{e}{h}\varphi_\alpha$ by φ_α they only occur in the two combinations: e^2/hc, a pure number (the so-called fine structure constant) and h/mc, an atomic length (the "wave length of the electron"). The meaning of the field equations differs radically from those of classical physics. It is not meant that the charge of the electron is really smeared over all space in accordance with (7), as Schrödinger held for a time, but rather that (7) is the probability of localization multiplied by the total charge $-e$. Before the equations can yield correct statistical predictions they must be subjected to the *process of quantization*, which is not entirely cleared up as yet. This process introduces the action constant h again. But that is obviously necessary if the theory is to be in a position to give an account of the atomistic constitution of matter, for in order to determine the three atomistic units of length, time, and mass we need in addition to the velocity of light c two further independent dimensional constants; and the field equations as they appear above contain but one, the length h/mc.

By this new situation, which introduces an atomic radius into the field equations themselves—but not until this step—my principle of *gauge-invariance*, with which I had hoped to relate gravitation and electricity, is robbed of its support. But it is now very agreeable to see that this principle has an equivalent in the quantum-theoretical field equations which is exactly like it in formal respects; the laws are invariant under the simultaneous replacement of ψ by $e^{i\lambda}\psi$, φ_α by $\varphi_\alpha - \frac{\partial \lambda}{\partial x_\alpha}$ where λ is an arbitrary real function of position and time. Also the relation of this property of invariance to the law of conservation of electricity remains exactly as before. The fact that the action integral remains unchanged by the infinitesimal variation

$$\delta\psi = i\lambda\psi, \qquad \delta\varphi_\alpha = -\frac{\partial\lambda}{\partial x_\alpha}$$

(λ an arbitrary function) signifies a dependence between the laws of matter and electromagnetism—which arise from the action integral by variation of the ψ and φ respectively. This identical relation consists in the fact that the law of conservation of electricity

$$\frac{\partial\rho_\alpha}{\partial x_\alpha} = 0$$

follows from the material as well as from the electromagnetic equations. The principle of gauge-invariance has the character of general relativity since it contains an arbitrary function λ, and can certainly only be understood in terms of it.

The tensor calculus is not the proper mathematical instrument to use in translating the quantum-theoretic equations of the electron over into the *general theory of relativity.* Vectors and terms are so constituted that the law which defines the transformation of their components from one Cartesian set of axes to another can be extended to the most general linear transformation, to an affine set of axes. That is not the case for quantity ψ, however; this kind of quantity belongs to a representation of the rotation group which cannot be extended to the affine group. Consequently we cannot introduce components of ψ relative to an arbitrary coordinate system in general relativity as we can for the electromagnetic potential and field strengths. We must rather describe the metric at a point P by local Cartesian axes $e(\alpha)$ instead of by the g_{pq}. The wave field has definite components $\psi_1^+, \psi_2^+; \psi_1^-, \psi_2^-$ relative to such axes, and we know how they transform on transition to any other Cartesian axes in P. The laws shall naturally be invariant under arbitrary rotation of the axes in P, and the axes at different points can be rotated independently of each other; they are in no way bound together. The formal aspects of our theory are similar to Einstein's recent attempts to unify electricity and gravitation; he too employs local Cartesian axes—*n*-legs, as he calls them—in place of the g_{pq}. But he assumes "*distant parallelism*," i.e., the axes in different points shall be so bound to one another that when rotated at one of them the axes in all other points automatically undergo the same rotation. I do not believe in this distant parallelism at all; there is no indication that Nature has availed herself of such an artificial geometry. I am convinced that if there is a physical content in Einstein's latest formal development it must come to light in the present connection. It now seems to me hopeless to seek a unification of gravitation and electricity without taking the material waves—the ψ field—into account.

For the purpose of analytic expression we need four *coordinates* x_p in addition to these local Cartesian axes. Let $e^p(\alpha)$ be the components of the vector $e(\alpha)$ in this coordinate system; the 4·4 quantities $e^p(\alpha)$ characterize *the gravitational field.* If $t(\alpha)$ be the components of a vector relative to the Cartesian axes, its contravariant components t^p relative to the coordinate system are given by

$$t^p = \sum_\alpha e^p(\alpha) t(\alpha).$$

Conversely the $t(\alpha)$ are obtained from the covariant components t_p, relative to the coordinate system by

$$t(\alpha) = \sum_p e^p(\alpha) t_p.$$

We consequently have upper and lower Latin indices, which refer to the coordinate system, and Greek indices, which belong to the axes. I have described how we are to juggle them.

Let $\int \mathfrak{H} dx$ be the action quantity of matter and the electromagnetic field ("matter in the extended sense"). We subject the $e^p(\alpha)$ to an arbitrary infinitesimal variation which vanishes outside a finite portion of the world and obtain an equation

$$\delta \int \mathfrak{H} dx = \int \mathfrak{t}_p(\alpha) \delta e^p(\alpha) dx.$$

It is here immaterial if or how the ψ and φ are varied, as the material and electromagnetic equations are assumed to hold; these latter state that the change in the action integral brought about by unrestricted variations of the ψ and φ vanishes. As we know, the general theory of relativity first enables us to give a general definition of the tensor density $\mathfrak{t}_p(\alpha)$ of energy by varying the metric field as above. The action integral is invariant under infinitesimal transformations of the coordinates; its variation consequently vanishes if the changes $\delta e^p(\alpha)$ are brought about by deformation of the coordinate system whereas the axes $e(\alpha)$ are held fast. This leads, as is well known, to the four components of the law of conservation of energy and linear momentum

$$\frac{\partial \mathfrak{t}_p{}^q}{\partial x^q} + \mathfrak{t}_q(\alpha) \frac{\partial e^q(\alpha)}{\partial x^p} = 0 \tag{8}$$

(This is admittedly not a true law of conservation in the general theory of relativity on account of the second term occurring in addition to the divergence.) Further, the action is invariant under an infinitesimal rotation, depending arbitrarily on position, of the local axes. This yields, as can be seen immediately, the law of symmetry of the energy tensor; it is here not identically fulfilled, but only in consequence of the material and electromagnetic equations. As we know, this symmetry is essentially identical with the law of conservation of angular momentum. Conservation theorems always result—and this is a general rule—from properties of invariance. We have also found an invariance property—gauge invariance—corresponding to the law of conservation of energy.

I must be more concrete. We wish to take over the principal term (4) of Dirac's action into general relativity. For this purpose we must know how a quantity of kind ψ is to be differentiated covariantly. Consider the fixed point $P=(x_p)$ and a neighboring point $P'=(x_p+dx_p)$; in both P and P' there is a set of axes $e(\alpha)$ and we denote the components of ψ referred to these by $\psi_p=\psi_p(P)$ and $\psi_p(P')$ respectively. The metric determines an infinitesimal parallel displacement which enables us to carry the axes $e(\alpha)$ in P over to P'; we thus obtain a set of Cartesian axes $e'(\alpha)$ in P', and we call the components of ψ in P' relative to these axes ψ'_p. The four differences $\delta\psi_p=\psi'_p-\psi_p$ depend only on the axes in P and transform with them in the same way as the ψ_p themselves. But $e'(\alpha)$ can be obtained from the axes $e(\alpha)$ in P' by an infinitesimal rotation do:

$$\delta e(\alpha)=e'(\alpha)-e(\alpha)=do(\alpha\beta)e(\beta)$$

and the ψ'_p arise from the $\psi_p(P')$ by the corresponding infinitesimal linear transformation dE. $\psi_p(P')-\psi_p(P)$ is the differential in the ordinary sense. Hence we finally have

$$\delta\psi=d\psi+dE\cdot\psi$$

dE depends linearly on the displacement $\overrightarrow{PP'}$ with the components $dx_p=(dx)^p$

$$dE=E_p(dx)^p.$$

Hence our formula yields

$$\psi_{(p)}=\frac{\partial\psi}{\partial x_p}+E_p\psi$$

as the components of the covariant derivative.

$$\bar{\psi}e^p(\alpha)S(\alpha)\psi_{(p)}$$

is an invariant which becomes a scalar density on division by the absolute value ϵ of the determinant $|e^{p}(\alpha)|$. Its integral must replace Dirac's quantity (4). After some difficulty we obtain as the action density of matter

$$(9) \qquad \frac{1}{i}\left\{\tilde{\psi}\,\mathrm{e}^{p}(\alpha)S(\alpha)\,\frac{\partial\psi}{\partial x^{\alpha}_{p}}+\tfrac{1}{2}\frac{\partial \mathrm{e}^{p}(\alpha)}{\partial x^{p}}\tilde{\psi}S(\alpha)\psi\right\}$$

$$+\frac{1}{\epsilon}f(\alpha)\psi S(\alpha)\psi.$$

$\mathrm{e}^{p}(\alpha)$ is the quotient $e^{p}(\alpha)/\epsilon$; $f(\alpha)$ is expressed in terms of the $e^{p}(\alpha)$ and their first derivatives in such a way that it vanishes with the derivatives.

On calculating the energy tensor belonging to this action quantity by the rules given above and then from it the total energy, the momentum and the moment of momentum by integration over the spacial cross-section $x_0=$const., we arrive back at the familiar assumptions of quantum theory. In particular the components of momentum are given by the spacial integral of

$$\frac{1}{i}\tilde{\psi}\frac{\partial\psi}{\partial x_{p}}(p=1,\ 2,\ 3)$$

and this is in accord with the basic association of the operators $\frac{1}{i}\tilde{\psi}\frac{\partial}{\partial x_1}$, with momentum p_1, as employed by Schrödinger. Also, the portion of the angular momentum due to the electron spin is not lacking, although one might at first be inclined to fear that by defining the components of angular momentum as the spacial integral of

$$x_2\mathrm{t}_3(0)-x_3\mathrm{t}_2(0),\ \cdots$$

only the orbital momentum would appear. This is naturally a powerful support for the assumption that (9) is actually the contribution to the action in so far as its dependence on the ψ is concerned.

Gravitation may be represented by the same action quantity as in Einstein's classical theory of gravitation. There then exists, as is well known a gravitational energy such that the total energy satisfies a true conservation law. If κ be the Einsteinian constant of gravitation the additive composition of the two terms in the action is accompanied by multiplying the gravitational contribution by $\frac{\kappa h}{c}$. This constant is the square of a length d. But d is much smaller than atomic dimensions—it is of the order 10^{-32} cm.

Now for the critical part, the electromagnetic field! Is it an appendage of gravitation or of the material field? A promising suggestion for the development of the first of these two viewpoints here offers itself. In the Dirac theory the influence of the electromagnetic potentials φ_p on matter is represented by the term

$$\varphi(\alpha)\tilde{\psi}S(\alpha)\psi. \tag{10}$$

Now we find that there is an additional term of exactly the same structure already in (9), without it being necessary to introduce a new entity into the theory in addition to gravitation and matter: we need only consider the $f(\alpha)$ which there appears, but which I did not write down explicitly, as the electromagnetic potential. This term arose from the gravitational potentials in such a way that, although it is invariant with respect to coordinate transformations, it is only invariant under rotation of the local axes in the restricted case in which the axis-systems in all points undergo the same rotation. Therefore, if we disregard the material field, there appears to be "distant parallelism" and we have a theory of exactly the same kind as Einstein's latest. But the calculations which I have made on this assumption seem to me to prove that it cannot be correct: we find no connection with Maxwell's equations, which are so well founded on observation; contrary to all experience we find that the potentials themselves, and not merely the field intensities, are of physical significance and our gauge invariance remains totally ununderstandable.

It is my firm conviction that we must seek the origin of the electromagnetic field in another direction. We have already mentioned that it is impossible to connect the transformations of the ψ in a unique manner with the rotations of the axis system; however we may attempt to accomplish this by means of invariants which can be used as constituents of an action quantity we always find that there remains an arbitrary "gauge factor" $e^{i\lambda}$. Hence the local axis-system does not determine the components of ψ uniquely, but only within such a factor of absolute magnitude 1. In the special theory of relativity, in which the axis system is not tied up to any particular point, this factor is a constant. But it is otherwise in the general theory of relativity when we remove the restriction binding the local axis-systems to each other; we cannot avoid allowing the gauge factor to depend arbitrarily on position. There then remains in the infinitesimal linear transformation dE of ψ, which corresponds to the given infinitesimal rotation do of the axis-systems, an arbitrary additive term $+id\varphi \cdot 1$. The complete determination of the covariant differential $\delta\psi$ of ψ requires that such a $d\varphi$ be given. But it must depend linearly on the displacement PP': $d\varphi = \varphi_p (dx)^p$, if $\delta\psi$ shall depend linearly on the displacement. On altering ψ by multiplying it by the gauge factor $e^{i\lambda}$ we must at the same time replace $d\varphi$ by $d\varphi - d\lambda$ as is immediately seen from this formula of the covariant differential. The principle of gauge invariance becomes self-evident. Our φ are the components of the electromagnetic potential, for they influence matter in the same way as these latter are known to do. Also conversely, they are influenced by matter in accordance with the law which experience has shown to hold for electricity.

$$f_{pq} = \frac{\partial \varphi_q}{\partial x_k} - \frac{\partial \varphi_p}{\partial x_q}$$

is indeed a gauge-invariant tensor. We build from it the familiar Maxwellian action, the analogue of the Dirichlet integral, the integrand of which we shall call $\mathfrak{D}(\varphi)$, and add it to the material term (9). The Maxwellian equations then result from varying the φ. Both terms in the action quantity have the same dimensions; the combining constant, a pure number, is the fine structure constant α. Our theory does not at present indicate any mathematical reason why this number has the numerical value which it actually has.

Until now we have stood, I believe, on solid ground; matter, electricity and gravitation are represented by the Dirac, the Maxwell and the classical Einstein action quantities, all of which have well withstood the test of observation (there could at most be a doubt concerning gravitation). The term (5) of the Dirac theory is, however, more doubtful. It must be admitted that if we retain it we can obtain all details of the line spectrum of the hydrogen atom—of one electron moving in the electrostatic field of a nucleus—in accord with what is known from experiment. But we obtain twice too much; if we replace the electron by a particle of the same mass and positive charge $+e$ (which admittedly does not exist in nature) the Dirac theory gives, contrary to all reason and experience, the same energy terms as for a negative electron, except for a change in sign. Obviously an essential change is here necessary. Furthermore, the theory in its present form contains only the *electron;* there can be no doubt that the *proton* must be introduced into the field equations before they are quantized. In place of one law of conservation of electricity we should have two, expressing the conservation of the number of electrons and protons separately. Our earthly physics sets before us the puzzle why an electron and a proton do not neutralize each other and release their total energy in the form of radiation —although astronomers, in the search for a more copious source of energy, play with the thought that this may actually occur in the stars.

The transformation of ψ under the influence of a rotation of the axis-system was described by the fact that the four quantities

$$\bar{\psi} S(\alpha) \psi$$

then transform among themselves as the four components of a vector fixed relative to the axes. But this description leaves open an even greater indeterminateness than was expressed by the gauge-factor $e^{i\lambda}$ above; the two pairs ψ^+ and ψ^- can indeed each be multiplied by an arbitrary number $e^{i\lambda^+}$, $e^{i\lambda^-}$; only then have we completely exhausted the arbitrariness inherent in the ψ. If we make use of it we find a double gauge invariance, correspondingly two laws of conservation of electricity and two electromagnetic potentials φ^+ and φ^- instead of one. The two components ψ_1^+, ψ_2^+ are to be ascribed to the proton, the remaining two ψ_1^-, ψ_2^- to the electron. It is not at all unpleasant that the wave quantity ψ of a single particle is again, as in the Pauli theory, reduced to the smaller ration of 2 components. But the term (5) involving the mass is now untenable, as it does not have the required gauge-invariance. The mass must be brought into the theory in some other way.

A possibility, which at first sight seems plausible, is to consider the gravitational term as the substitute for the mass. For mass is a gravitational effect; it is the flux of the gravitational field through a surface enclosing the particle in the same sense that charge is the flux of electric field. In a satisfactory theory it should be as impossible to introduce a non-vanishing mass without gravitational field as it already is to introduce charge without electromagnetic field. And the gravitational term actually introduces a constant d of the dimensions of length; but unfortunately it lies far under the order of magnitude of the wave length of the electron $\frac{h}{mc} \backsim 10^{-10}$ cm. This is the old dilemma—that the constant of gravitation falls so far out of the range of the other constants of nature. For this reason it seems impossible that gravitation could help us out. I suggest the following way: Be bold enough to leave the term involving mass entirely out of the field equations. But the integral of the total energy density over space yields an invariant, and at the same time constant, mass; *require of it that its value be an absolute constant of nature* m which cannot vary in value from case to case. This introduction of mass is born of the idea that the inertia of matter is due to its energy content.

The two electromagnetic fields may excite some doubt, but as far as I can see one can only thus obtain two conservation theorems, and from the group-theoretic standpoint their appearance seems almost unavoidable. In place of the Maxwellian $\mathfrak{D}(\varphi)$ we should introduce the electromagnetic action density

$$a\mathfrak{D}(\varphi^+)+2b\mathfrak{D}(\varphi^+, \varphi^-)+a\mathfrak{D}(\varphi^-).$$

The middle term is the symmetric bilinear form corresponding to the quadratic $\mathfrak{D}(\varphi)$. This quadratic form with coefficients a, b, a must be definite, $a^2-b^2>0$, if the energy of radiation is to be always positive; one wishes indeed to retain this property, for according to the testimony of experience a field of light has a lowest energy level—darkness. The middle term is responsible for the interaction between protons and electrons. To the same approximation in which the electromagnetic energy arising from this middle term and the gravitational energy can be neglected in comparison with the remaining energy, the mass m introduced above is broken up into two parts m^+ and m^-, the masses of the proton and the electron, which contain only ψ^+, φ^+; ψ^-, φ^- respectively. The masses of the proton and the electron will consequently not be exactly constant but will vary somewhat from quantum state to quantum state. Mass already behaves in this way in the Dirac theory; Dirac's constant m is not the inertial mass of the electron, but rather a "mass factor," a number which occurs as a common factor in the inertial masses of the electron in various quantum states.

I believe that these ideas point in the direction of the future development of quantum theory; on the other hand, I am not certain that the above sketch is correct in every point. I have not as yet found a satisfactory approach to that approximation which reduces the problem of a hydrogen atom to a linear one by neglecting the field of radiation and its reaction on the atom in comparison with that portion of the electromagnetic field, known *a priori*, which binds the proton and the electron to one another in accordance with the Coulomb law.

Another difficulty which stands in the way of a comparison with experience is that the field equations must first be quantized before they can be applied as a basis for the statistics of quantum transitions. But our theory is also hopeful in this respect inasmuch as the anti-symmetric Fermi statistics of the electrons, corresponding to the Pauli exclusion principle, here necessarily leads to the symmetric Bose-Einstein statistics of photons.

Hermann Weyl.

R·R·DONNELLEY & SONS·COMPANY·THE·LAKESIDE·PRESS·R·CHICAGO·

CONTENTS

Vol. XVI January, 1929 No. 1

PAGE

THE AIMS AND METHOD OF HIGHER LITERARY STUDIES—Three public lectures delivered in the Physics Amphitheatre of the Rice Institute, January 3, 4, and 5, 1929, by LOUIS CAZAMIAN, Professor of Modern English Literature and Civilization in the Faculty of Letters of the University of Paris.

I. The Lure of Sources 1
II. The Object of Criticism 17
III. Eugenius; or, Everybody His Own Critic 32

Vol. XVI April, 1929 No. 2

MODERN SPAIN AND ITS LITERATURE—A course of three lectures delivered at the Rice Institute on March 20, 21, and 22, 1929, by WALTER FITZWILLIAM STARKIE, Litt.D., Professor of Spanish in Trinity College, Dublin.

I. Personalities of Modern Spain 47
II. The Drama of Spain 66
III. A Modern Don Quixote 86

Vol. XVI July, 1929 No. 3

THE PHI BETA KAPPA IDEAL—An address delivered by HENRY OSBORN TAYLOR, D.Litt., L.H.D., of New York City, at the installation of the Beta of Texas Chapter of the Phi Beta Kappa Society at the Rice Institute, March 2, 1929 . 111

MODERN MUSIC—Lecture-Recital by ARTHUR HONEGGER, of Paris, under the auspices of the Rice Institute Lectureship in Music, Houston, March 6, 1929.
I. Programme of the Lecture-Recital 122
II. Lecture on Modern Music 123

THE SUPREME VALUES IN HUMAN NATURE—A sermon preached to the fourteenth graduating class of the Rice Institute, by the Reverend WILLIAM DOUGLAS MACKENZIE, D.D., LL.D., President of the Hartford Seminary Foundation, Hartford, Connecticut 131

THE HABIT OF CRIME IN THE UNITED STATES—An address delivered at the fourteenth commencement convocation of the Rice Institute by WILLIAM EDWARD DODD, Ph.D., LL.D., Professor of American History at the University of Chicago 142

The Rice Institute Pamphlet is published at Houston, Texas, and is issued in January, April, July, and October. It was entered as second-class matter, April 15, 1915, at the post-office at Houston, Texas, under the Act of August 24, 1912